



Horizon 2020 Program

Dynamic countering of cyber-attacks

SU-ICT-2018



Cyber security 4.0: Protecting the Industrial Internet of Things

D1.1: C4IIoT Innovations for Industrial IoT Systems[†]

Abstract: In this deliverable we briefly summarize significant new research results and new commercial products that have been published or released on the market, since the submission of the project proposal in 30/08/2018, in the technological areas where C4IIoT aims to advance the State-of-the-Art (SotA thereafter). The goal of this document is to insure that the starting point for these targeted advances is up to date. It thus contains one section for each technological area listed as a target for advance in section 1.4.1. of the proposal which describes the project's offerings beyond the SotA.

[†] *The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833828.*

Contractual Date of Delivery	30/09/2019
Actual Date of Delivery	30/09/2019
Deliverable Security Class	Public
Editor	<i>Jacques Robin, UPIPS</i>
Contributors	<i>Dragan Danilovic, VIP</i>
	<i>Alberto Terzi, HPE</i>
	<i>Marie-Noëlle Lepareux, TGS</i>
	<i>Konstantinos Fysarakis, STS</i>
	<i>Omri Soceanu, IBM</i>
	<i>Jacques Robin, UPIPS</i>
	<i>Antonio Escobar, Ji Zheng IFAG</i>
	<i>George Bravos, ITML</i>
	<i>Dusan Jakovetic, Srdjan Skrbic, Milan Lukic, Ivan Mezei, Dejan Vukobratovic, Milos Savic UNSPMF</i>
	<i>Ilias Spais, AEGIS</i>
Quality Assurance	<i>Georgia Sakellari UoG</i>
	<i>Giorgios Vasiliadis FORTH</i> <i>Dusan Jakovetic, UNSPMF</i>

The C4IIoT Consortium

FOUNDATION FOR RESEARCH AND TECHNOLOGY HELLAS (FORTH)	Coordinator	EL
CENTRO RICERCHIE FIAT SCPA (CRF)	Principal Contractor	IT
INFINEON TECHNOLOGIES AG (IFAG)	Principal Contractor	DE
THALES SIX GTS FRANCE SAS (TGS)	Principal Contractor	FR
HEWLETT PACKARD ITALIANA SRL (HPE)	Principal Contractor	IT
COMMISSARIAT A L'ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES (CEA)	Principal Contractor	FR
IBM ISRAEL - SCIENCE AND TECHNOLOGY LTD (IBM)	Principal Contractor	IL
AEGIS IT RESEARCH UG (AEGIS)	Principal Contractor	DE
UNIVERSITE PARIS I PANTHEON-SORBONNE (UPIPS)	Principal Contractor	FR
INFORMATION TECHNOLOGY FOR MARKET LEADERSHIP (ITML)	Principal Contractor	EL
SPHYNX TECHNOLOGY SOLUTIONS AG (STS)	Principal Contractor	CH
UNIVERSITY OF NOVI SAD FACULTY OF SCIENCES (UNSPMF)	Principal Contractor	SR
UNIVERSITY OF GREENWICH (UOG)	Principal Contractor	UK
VIP MOBILE D.O.O.(VIP)	Principal Contractor	SR

Document Revisions & Quality Assurance

Internal Reviewers

1. *Giorgios Vasiliadis, FORTH*
2. *Dragana Bajovic, Dusan Jakovetic, UNS*

Revisions

Version	Date	By	Overview
C4IIoT.D1.1.2	30/09/2019	Editor	Final document.
C4IIoT.D1.1.3	22/09/2019	#1, #2	Comments on first draft
C4IIoT.D1.1.2	13/09/2019	Editor	First draft.
C4IIoT.D1.1.1	03/08/2019	#2	Comments on the ToC.
C4IIoT.D1.1.0	13/07/2019	Editor	ToC.

Table of Contents

LIST OF TABLES.....	6
LIST OF FIGURES.....	7
LIST OF ABBREVIATIONS	8
EXECUTIVE SUMMARY	9
1 INTRODUCTION	10
1.1 LEVEL-1 SECURITY: HARDWARE-ENABLED SECURITY	11
1.2 LEVEL-2 SECURITY: SECURITY ENABLED BY HORIZONTAL DEVICE-TO-DEVICE COMMUNICATION	13
1.3 LEVEL-3 SECURITY: SECURITY ENABLED BY MACHINE LEARNING-BASED BEHAVIOURAL ANALYSIS AND COGNITIVE SECURITY CAPABILITIES	13
1.4 MITIGATION ENGINE.....	14
1.5 NOMINAL DATA USERS: SITUATIONAL AWARENESS VIA INFORMATIVE VISUALIZATIONS	14
1.6 DATA FEDERATION PARTNERS: PRIVACY-AWARE ANALYTICS AND ACCOUNTABLE DATA PROCESSING ...	15
2 SECURED AND TRUSTED ENVIRONMENTS	16
2.1 PROPOSAL'S STATED OFFERING BEYOND AUGUST 2018.....	16
2.2 OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018	16
2.3 OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	18
3 ENCRYPTED TRAFFIC ANALYSIS	19
3.1 PROPOSAL'S STATED OFFERING BEYOND AUGUST 2018 SOTA	19
3.2 OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018	19
3.3 OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	20
4 VULNERABILITY ANALYSIS AND MITIGATION.....	22
4.1 PROPOSAL'S STATED OFFERING BEYOND AUGUST 2018 SOTA	22
4.2 OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018	23
4.3 OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	23
5 DECENTRALIZED SELECTIVE ACCESS MANAGEMENT.....	24
5.1 PROPOSAL'S STATED OFFERING BEYOND AUGUST 2018 SOTA	24
5.2 OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018	24
5.3 OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	25
6 DYNAMIC VERIFICATION AND RE-CONFIGURATION OF SELF-ADAPTIVE SYSTEMS ...	26
6.1 PROPOSAL'S STATED OFFERING BEYOND AUGUST 2018 SOTA	26
6.2 OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018	27
6.3 OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	29
7 SECURE-BY-DESIGN IIOT DEVICE FABRICATION	30
7.1 PROPOSAL'S STATED OFFERING BEYOND AUGUST 2018 SOTA	30
7.2 OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018	31
7.3 OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	31
8 MACHINE LEARNING AT THE EDGE.....	33
8.1 PROPOSAL'S STATED OFFERING BEYOND AUGUST 2018 SOTA	33
8.2 OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018	34
8.3 OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	36
9 SECURITY-AWARE OFFLOADING DECISION SUPPORT (UOG)	39
9.1 PROPOSAL'S STATED OFFERING BEYOND AUGUST 2018 SOTA	39
9.2 OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018	39
9.3 OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	40
10 FORENSICS VISUALIZATION	41

10.1	PROPOSAL’S STATED OFFERING BEYOND AUGUST 2018 SOTA.....	41
10.2	OVERVIEW OF RESEARCH ADVANCES SINCE AUGUST 2018.....	41
10.3	OVERVIEW OF INNOVATIVE COMMERCIAL PRODUCTS RELEASED SINCE AUGUST 2018.....	42
11	REFERENCES	43

List of Tables

Table 1: Popular ML/DL tools	36
------------------------------------	----

List of Figures

Figure 1: C4IIoT conceptual Framework.....	11
Figure 2. C4IIoT Continuous Verification	22

List of Abbreviations

AOP	Aspect-Oriented Programing
API	Application Programming Interface
C4IIoT	Cybersecurity for the Industrial Internet-of-Things
C&C	Command and Control
DPI	Deep Packet Inspection
DSPL	Dynamic Software-intensive systems Product Line
DSPLE	DSPL Engineering
EC	European Commission
FragOP	Fragment Oriented Programming
GORE	Goal-Oriented Requirement Engineering
IIoT	The Industrial Internet-of-Things
IoT	The Internet-of-Things
M@RT	Models at Run-Time
MDE	Model-Driven Engineering
MILP	Mixed Integer Linear Programming
NB IOT	Narrowband IoT
SAS	Self-Adaptive Systems
SASE	SAS Engineering
SMT	Satisfiability Modulo Theory
SPL	Software-intensive systems Product Line
SPLE	SPL Engineering
SotA	State-of-the-Art
WP	Work Package

Executive Summary

In this deliverable we briefly summarize significant new research results and new commercial products that have been published or released on the market, since the submission of the project proposal in 30/08/2018, in the technological areas where C4IIoT aims to advance the State-of-the-Art (SotA thereafter). The goal of this document is to insure that the starting point for these targeted advances is up to date. It thus contains one section for each technological area listed as a target for advance in section 1.4.1. of the proposal which describes the project's offerings beyond the SotA.

1 Introduction

Industrial IoT collects and analyzes data to deliver insights that help industrial companies become more agile, making better informed business decisions more quickly than ever before. This means better quality control, and more efficient, streamlined supply chain management. It also benefits predictive maintenance, field service, energy and facilities management, and asset tracking.

In the age of digitized everything, security breaches and hacker attacks are no longer even newsworthy. The spread of cloud services and the advent of the Internet of Things have urged enterprises to enhance security and rethink their company policies. The overall complexity of a smart factory IoT system is extensive, and the number of security loopholes subsequently increases to a dramatic extent. Clearly, traditional firewalls and antiviral systems will not be sufficient; the complex IIoT infrastructure demands something more advanced. An IIoT network needs an advanced security system, not only to ensure a non-disruptive smart factory workflow, protect employees and assets, but also to secure business-critical information from competitors.

In order to meet the challenges mentioned above, C4IIoT will design, build and demonstrate a novel and unified Cybersecurity 4.0 framework for malicious and anomalous behaviour anticipation, detection, tracking, mitigation, and end user informing. The framework develops novel security-enabling components and integrates them in innovative ways to provide a holistic and disruptive security-enabling solution for minimizing the attack surfaces in IIoT systems. Our solution exploits emerging security software and hardware protection mechanisms (including trusted and verifiable computation systems and environments and advanced verification tools), state of the art machine and deep learning and analytics, including privacy-aware analytics, novel encrypted network flow analysis, secure-by-design IIoT device fabrication, and blockchain technologies, to provide a viable scheme for enabling security and accountability, preserve privacy, enable reliability and assure trustworthiness within evolving IIoT applications. This will be achieved by a novel cybersecurity mechanism that is carefully orchestrated across all infrastructure elements involved within an IIoT system (e.g., IIoT devices, field gateways, cloud resources) and is based upon analysis of various data flows (e.g., IIoT device data, encrypted network flows).

Our framework includes layered organization according to the physical hardware and key infrastructure block. C4IIOT framework is organized into three layers: the edge nodes layer, the field gateways layer, and the cloud layer.

- The *edge nodes* layer includes the devices and sensors that generate the data that feed the whole framework. Standard commercially available IoT devices will be used based on different IIoT connectivity options, ranging from low-power Wi-Fi based solutions to cellular-based solutions (NB IOT, LoRa, LTE-M).
- All the IIoT devices are connected to *field gateways* which form the second layer of our architecture. The field gateway might be under the control of the device owner (for example it might IoT gateway as part of the company's private network) or not (for example in the case of smart meter infrastructures where the concentrator is away from the homes). As an intermediate layer, the field gateway acts as an element with increased computational power to perform security-related (e.g. anomaly detection) actions with significantly lower latencies than the cloud since it resides close to the edge nodes. The role of the field gateway in C4IIOT is dual. Besides acting as a

Software-Defined-Networking (SDN)-enabled network node it will also include a local offloading/outsourcing decision mechanism.

- The *cloud layer* is the main data handling component of the architecture. In addition to consuming data and providing the appropriate services to users, the cloud has an essential role for cybersecurity. It communicates with the field gateways through a cloud gateway, responsible for the communication functionalities. It hosts several services, including support of secure data streams, as well as services that ensure the authenticity and integrity of the data flows, the Command and Control (C&C) and the maintenance. The security services are realized through the two main modules Behavioural analysis and cognitive security module, and the Mitigation engine.

Here we are introducing cybersecurity framework which will be built upon techniques and methodologies to enable strong authentication and authorization models, code signature and verification models, secure execution, compliance, accountability, behavioural analysis, and privacy-preserving multi-party analytics.

The innovative framework enables forecasting and detecting threats at different levels: i) hardware-enabled security; ii) security enabled by horizontal device-to-device communication through distributed ledger technologies; and iii) security enabled by context-aware intelligence for detecting anomalous or malicious behaviour. It also provides a comprehensive solution for mitigation, informing end users and activating data federation partners. **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.** Figure 1 below illustrates the C4IIoT conceptual framework which we first briefly describe in introductory part, and then provide detailed description in the following subsections.

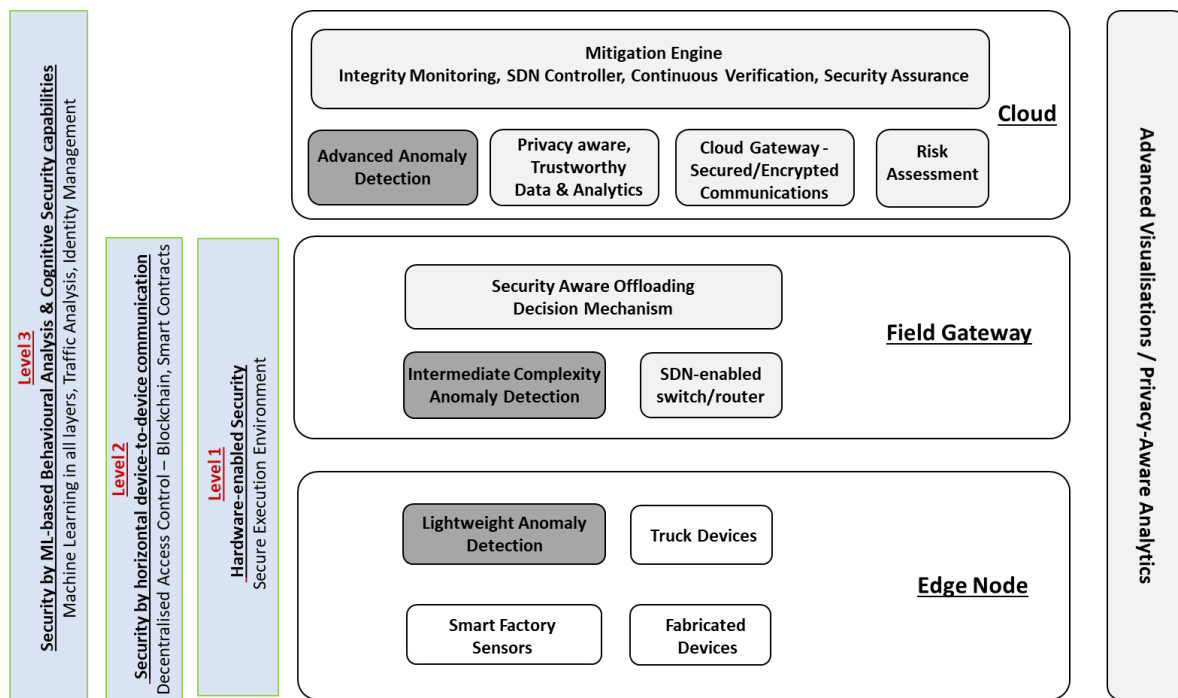


Figure 1: C4IIoT conceptual Framework

1.1 Level-1 security: Hardware-enabled Security

Hardware plays a critical role in today's security landscape. The increasing number of vulnerability disclosures requires a more rigorous approach to secure hardware designs and to

implement secure mechanisms. In order to meet strict security standards, C4IIoT will use the mechanisms that rely on: i) isolated execution environments; ii) trusted execution; iii) remote attestation; and iv) robust access control.

A code integrity mechanism verifies the integrity of the code running within the trusted hardware enclaves which are areas of execution in memory with more security protection and the kernel code as well. Enclaves have hardware-assisted confidentiality and integrity-added protections to help prevent access from processes at higher privilege levels. The code in the trusted enclave is responsible for ensuring the integrity of the kernel code at runtime, and detecting any modifications by malicious software, once the system is running. At the same time, data integrity mechanisms will protect sensitive and mission-critical data. A special emphasis will be put on targeting data leaks, control-flow hijacking, and API usage through techniques ranging from deep semantic bug finding to full verification depending on the requirements from the end-users.

The foundation of the proposed security strategy leverages the capabilities of secure elements that lie in the hardware chip. To ensure that the three main concepts of cyber security (confidentiality, integrity and authentication) are implemented in a proper way, it is necessary to utilize strong and up-to-date cryptographic algorithms. Momentarily, Advanced Encryption Standard (AES) is used as a de-facto standard for symmetrical encryption of payload data between two end points, with key lengths of 128 or 256 bits depending on security requirements. Keys can be pre-shared between the end device and cloud instance and used over the whole device lifetime, but a more secure strategy is to use session keys with predetermined expiration period, using public-key cryptography algorithm for key exchange sessions. Elliptic-curve cryptography (ECC) is a logical choice for small footprint public-key implementations, providing the same degree of security as RSA with significantly shorter key size. Finally, for secure hashing, it is recommended to use function belonging to SHA-2 family with an appropriate digest size.

It is recommended to use a hardware element with secured non-volatile memory for key storage, a random number generator with sufficient degree of entropy, and an embedded coprocessor to handle the utilization of cryptographic algorithms. This way, the device firmware can be made oblivious of the security keys and therefore less susceptible to hacking. It is necessary to use a mechanism to prevent repeat attacks. In that regard, cryptographic nonce and/or sequence counters can be appended to the payload data, and duplicate packets are to be dropped on the server side.

The hardware-enabled components will further allow secure access to every layer of the architecture (edge nodes, field gateways, cloud), facilitating strong authentication and authorization schemes that are based on single security tokens. These security tokens, stored in hardware-enforced secure enclaves will be responsible for identity and access management, ensuring accountability and preventing false identities to commit fraud, acquire sensitive data, commit data theft, or manipulate the system by any means. The described functionalities constitute overall a Secure execution environment that is built upon several partners' technologies, including FORTH's privacy-enhanced execution, IFAG's OPTIGA Trust X device for authentication and encryption, IFAG's OPTIGA TPM for secure booting and remote attestation, TSG's access control infrastructure.

1.2 Level-2 security: Security enabled by horizontal device-to-device communication

This level of security corresponds to a decentralized approach, with intermediate, field gateways layer that aggregates IIoT edge node groups information, and ensure reduced reaction times. Decentralized communication between IIoT devices shall be enabled, without a centralized communication service between them, and field gateway, as intermediate layer will have increased computational power to perform security-related actions with significantly lower latencies than the cloud, since it resides close to the edge nodes. Field gateway will also implement security offloading decision mechanism, which evaluates the trade-off between confidence in anomaly detection and cost of offloading to the cloud (high computational power, high latency), and automatically triggers appropriate option. This will allow running actions closer to the edge device, whenever possible, and minimize communication to the cloud. The blockchain technology will also be applied, and it will form a trusted ledger of transactions to enhance system's accountability, auditability and reliability. Additionally, blockchain solution will allow for processing of transactions and validating produced data. However, decentralized solutions built on blockchain technologies do not scale yet to the amount of data aggregated by IIoT devices. Moreover, in case a privacy policy or a business decision changes, the system needs to be able to easily revoke access to data. Henceforth, the IIoT aggregated data and analytics results will be stored on a restricted secure storage on the cloud, after encrypting the data, and only the link to the restricted storage will be recorded by the implemented blockchain based solution. When a privacy policy or a business decision changes, the decentralized access management solution will revoke the access to the restricted storage to some data federation partners as required. In addition to enforcement of policies, the records documented on blockchain can be also used to prove compliance with defined centralized rules and for forensics purposes. Level-2 security of the C4IIoT framework will be built upon multiple technologies brought by the partners, including IFAG's smart contracts management module and IBM's Data Policy and Consent Management solution. IBM's DPCM is a fully centralized solution, so in C4IIoT IBM will provide a decentralized privacy aware data access management solution, as described in Section 5.

1.3 Level-3 security: Security enabled by machine learning-based Behavioural Analysis and Cognitive Security capabilities

Context aware intelligence for detecting anomalous or malicious behaviour is needed because even strong hardware protections (level-1) built in a device, the code protections, and secure communication protocols could be bypassed or exploited. Furthermore, device-to-device communication (level-2 security) may have a localized effect (e.g., within an IIoT device geographical vicinity), and may not cover an IIoT system in its entirety. To address risks identified on level-1 and level-2 and gain deeper understanding of the whole environment, a separate mitigation mechanism is required. This mechanism will be based on contextual information and context will be built by using different inputs collected from the edge nodes and field gateways, and also regarding IIoT system as a whole, not considering devices in isolation. Such inputs could consist of regular sensor output values (e.g., GPS coordinates), monitoring data, and multiple sensors output that could be used to convey complex information (for instance, change in the rotational speed of a motor). Using these inputs, models will be trained to detect advanced threats and anomalies. Anomaly detection is the first step in raising alarms and devising plans for compensating and responding to the anomaly and to enable it normal behaviour of the system has to be determined regarding all relevant inputs. A comprehensive IIoT solution in the security domain has to balance

between reacting as quickly as possible (time criticality) and the achieved accuracy of detection. This could be achieved by implementing parallel and distributed instances of ML algorithms across all layers (edge nodes, field gateways, cloud) and coupling with decisions provided by the security offloading mechanism.

Security offloading mechanism is an innovative and carefully designed solution that delegates different security tasks to different layers of the architecture. From the ML algorithmic perspective data partitioning is performed across all edge devices' data to update the global model stored in the cloud in accordance to federated learning concept. Security offload technologies are still in the early stages and typically geared towards manual decision. Additional security mechanism will be provided by focusing on encrypted network data flow analysis since current industry wide solutions based on deep packet inspection (DPI) are being phased out. Besides requiring very high processing power and specialized hardware, DPI algorithms are processing non-encrypted payloads that are not suitable for secure solutions. Encrypted communication network traffic flow analysis relies on detecting specific patterns that are detected by using machine and deep learning techniques applied to communications meta-data that can be found inside packet headers or times, including packet lengths, timestamps, directions and inter-arrival times. This approach is already being used to identify web pages transferred over encrypted tunnels or fine-grained application-level events (*e.g.*, messaging) and will further strengthen the IIoT data flows trustworthiness by capturing attacks or anomalies inside the real network traffic. The outcome of this step will provide insights on the proper metadata handling and processing in order to produce network signatures that will be used in order to enrich current network inspection systems' functionality. In the training stages, ground-truth datasets of malicious traffic from industrial partners use cases will be collected and analysed to derive signatures based on traffic metadata.

1.4 Mitigation Engine

The Mitigation engine performs mitigation measures, based on the inputs provided by the behavioural analysis and cognitive security framework. Actions taken by the mitigation engine affect several layers of the architecture: (i) the edge nodes layer, where actions such as FW/SW updates, and reconfiguration of the edge nodes, are taken on behalf of the user; (ii) the field gateways layer, where traffic containing suspicious signatures/patterns is filtered using centralized rules and blacklists, thereby protecting the edge nodes from harmful data (SW/FW data, application malware, suspicious code ...); (iii) cloud layer, where secure policies, access control, and trusted execution are applied. The communication between the mitigation engine, and the other infrastructure constituents of the C4IIoT platform (edge nodes, field gateways, cloud infrastructure) is accomplished through the C&C and maintenance signalling. The mitigation functionality of the Mitigation engine will be built upon multiple partners' technologies, including the UP1PS VariaMos model-based self-adaptive service reconfiguration search, TSG's SDN controller for traffic filtering, and CEA's BINSEC program analyzer for continuous verification, and the STS's security assurance and dynamic certification platform.

1.5 Nominal data users: situational awareness via informative visualizations

In order to have a solution with strong security properties and high level of privacy that utilize situational awareness an advanced informative visualisation solution will be offered within

the C4IIoT framework. It can inform users on their security/privacy levels, while providing warnings and assisting them in handling security and privacy related incidents.

Although having advanced operating infrastructure, it will feature an easy to use and understandable way to visually communicate to the end users to help them understand the benefits of offered features without necessity to understand the complexity of the underlying technologies. The interface elements will be based on intuitive and flexible design that can be configured and used with any device. End users will be able to manage the monitoring of the infrastructure, look into the details and insights of the incidents that occurred and browse generated statistics.

The functionality that features secured and privacy-preserving stored data, analytical operations with advanced visualisations that can foster forensics analysis will be based upon the AEGIS' Advanced Visualization Forensics tool. It also features temporal inspection of given metrics and combinations of relevant measurements in interactive visual representations, facilitate investigation on security alerts and help users to better protect against future attacks and vulnerabilities. The inputs to the tool will be taken from all appropriate architecture layers within the C4IIoT framework.

1.6 Data federation partners: Privacy-aware analytics and accountable data processing

Seamless integration and ingestion of heterogeneous data and adoption of collaborative analytics, without exposing private or sensitive information, is fundamental for future secure collaborative environments. To achieve this goal, the system should allow the user to select which data (or portion of the data) will be selected in collaborative tasks, that will protect data and selectively allow access according to user defined criteria. A number of privacy-preserving storage and analytics technologies are emerging, leading to the concepts of federated or collaborative data analytics, in which end users engage in controlled and privacy-aware exchange of information for the benefit of performance of all involved actors. Privacy aware machine learning and decentralised access management will be key to achieving the goal of privacy preservation and accountability of the framework. Regarding privacy-aware analytics, cryptography-based approaches, and non-cryptography-based approaches like differential privacy will be considered. In addition, federated learning methods in which, instead of exchanging end user data, appropriate descriptions of machine learning models are exchanged, are becoming increasingly efficient. Moreover, the project will leverage secure hardware (such as Intel SGX) in order to ensure the compliance of the requested data access criteria.

2 Secured and trusted environments

2.1 Proposal's stated offering beyond August 2018

Following the description of security in the proposal WP3, the project will ensure the provision and configuration of infrastructure resources through efficient resource management and orchestration in a secured and trusted environment.

To develop the C4IIoT core of Level-3 security mechanism which consists of the development of behavioral models that will enable the analysis of the behavior of multiple IoT devices; to develop mitigation and immune reaction mechanisms across different layers; To design and develop the building components composing the C4IIoT trust infrastructure.

2.2 Overview of research advances since August 2018

In the C4IIoT context, with respect to the discussions that led to the design of the initial solution, there were no major developments. The major developments concern the awareness of the security problem in general and in the field of operation technology and industrial internet of things area. Let us add an introduction to explain the several point:

There are several aspects to consider regarding the security of industrial equipment and in particular for the IIoT (IIoT Industrial Internet of things) and IOT.

First, we need to set the scene: we are in the Cyberworld where live more than 40 billion of elements that continually grow and Cybercrime, Cyberterrorism, Cyberwarfare, Cyber Espionage hacktivism and general cyber threats are the “bad guys”.

Considering that everything from refrigerators to sprinkler systems, form tower crane to supply assembly line are wired/wireless interconnected, and while these devices have made life easier, they have created new attack vectors for hackers.

IoT elements are poised to become more pervasive in our lives than mobile phones and will have access to the most sensitive personal data such as social security numbers and banking information. As the number of connected IoT devices constantly increase, security concerns are also exponentially multiplied. A couple of security concerns on a single device such as a mobile phone can quickly turn to 50 or 60 concerns when considering multiple IoT devices in an interconnected home or business. Based on the importance of what IoT devices have access to, it is important to understand their security risk.

In this Cyberworld the 90% of devices collect personal information, 70% utilize unencrypted communication and the attackers are generally strongly motivated to reach the goal.

We identify that in the first three quarter of 2019, the cybercrime maintain the leadership, and the cyber warfare and activism grow up. In this scenario, the main attack vector is the malware, and the human factor, that cause incident. Specifically for the Industrial environment, the communication has poor security (the remote controller of a supply assembly line the security based on embedded password and a simple sniffer could interfere in the wireless communication). The data is stored in legacy database historian (as example), and poor human interfaces, based on HMI/PLC or keyboards that with a specific keys sequence allow the access.

In the Cyberworld we described several risk, but what are the risk in the traditional world?

The report of the world economic forum [1] identifies five areas: economic, geopolitical, environmental, societal, and technological; for each area, risks are identified. Cyber-attacks are the first technology risk, in term of likelihood they represent the fifth risk, and in term of impact the seventh risk.

That it means this immaterial world could have a huge impact in the real world.

The threat landscape based on ENISA report [2] provides some input: the European Union Agency for Network and Information Security provide this report of the security threat trend.

The malwares are the dominating attack vector, the target are mainly the endpoint then servers, cloud, IoT, mobile devices, Industrial Control System (ICS). Interesting that the last year the crypto jacking (a program that secretly mines cryptocurrency) increased.

What we know, and is an objective practice of our economic model, is the continuous search for high-performance and low-cost equipment that meets basic functions without any further focus on security.

What did we discover from the analysed IoT devices from manufacturers of TVs, webcams, home thermostats, remote power outlets, sprinkler controllers, hubs for controlling multiple devices, door locks, home, alarms, scales, and garage door openers, industrial sensor and industrial component?

We discovered the following architectural component:

- A majority of devices includes some form of cloud service.
- Almost devices include mobile applications that can be used to access or control the devices remotely.

From the current world we know that:

- Security is always a trade-off – a balance between cost risk and benefit
- Security is a process, not a tool – is useful to have the best-armoured door left open
- Complex systems introduce a wider attack surface
- Interconnected devices extend the vulnerability footprint
- A single weak element can affect the overall security architecture
- Humans are smart, creative and continually adopt new tactics and techniques to attack
- Humans are the weakest link in a security chain
- Patching as a defence needs to be readdressed (low cost devices)

In addition, concerning attackers we know that:

- We live in a world where the security threats are more complex & sophisticated
- They can be external or internal or they can represent malicious or unintentional actions.
- The criminal marketplace are full of solutions, and highly motivated to gain access to information for profit, politics & corporate espionage

We also know that complex regulatory (e.g., the GDPR [3]), and industry specific issues continue to increase.

Not only do we have to protect, but also we must be able to adapt rapidly to this every changing landscape of pressures while responding to the agility demands of the business.

Security at the edge does not protect the core. In addition, moving to the cloud without the right principles and processes creates new vulnerabilities and exposure that lead to attacks and

potential failure. Enterprises need to build security into every layer of the stack: from silicon to apps, and everywhere, from edge to cloud.

Security should not be an afterthought, or something you bolt on to your infrastructure; the costs are too great. A competitive organization will employ a secure continuum from intelligent edge to enterprise core, whether on premise or cloud, so that they are guarded against disruption.

Digital enterprise security requires a holistic approach, also because the devices are more intelligent and smart including the new buzzword IoT/IIoT

In addition to what we have described of potential attacks we have a situation of side attacks that affected the chain of the device, just remember the Meltdown and Spectre [4] hardware vulnerability affecting microprocessors.

How to prevent or mitigate an attack on the CPU? An option to prevent and mitigate the attack is to have a validation of the code to execute. To explain, the first code executed in a device arrives from the Boot ROMs. As indicated previously, the very first code executed by a CPU is fetched from a non-volatile NOR flash which is known as Boot ROMs. So the Boot Rom must be protected, but how protecting Boot ROMs from malicious attacks and also detecting intruders and recover from such attacks with the end objective of achieving resiliency? A solution could be other than the TPM the solution identified by Hewlett Packard Enterprise: HPE has been the first to introduce Silicon Root of Trust [5], which establishes trust by validating the most critical segment of boot code which resides inside a SPI NOR flash. Any attempt to alter the flash contents or to damage it must be detected and remedial action should be initiated, this solution creates a secure and trusted execution of the code.

To conclude we need to find methods for Secure Trust and Secure Execution, that is valid like the SRT or the TPM for the IOT device, generally the TPM model is a valid model for protecting cryptographic keys,

Trusted Platform Module is an international standard for a secure crypto-processor, a dedicated microcontroller designed to secure hardware through integrated cryptographic keys. The primary scope of TPM is to assure the integrity of a platform

That allows a secure trust to be carried out with appropriate checks, for secure execution the matter is more complex as it concerns an area where you can safely execute a code, in the IOT area you can assimilate these areas to the device emphasizing the concept of secure trust

2.3 Overview of innovative commercial products released since August 2018

There are no significant updates about Secured and trusted environments since August 2018 in commercial products, as old HPE has been the first to introduce Silicon Root of Trust, other hardware players introduced a similar technology.

3 Encrypted traffic analysis

3.1 Proposal's stated offering beyond August 2018 SotA

Following the description of the level-3 security in the proposal, the project will specify an offering of encrypted network analysis aiming to identify a malicious behavior hidden in encrypted streams. The aimed analysis will be supported by ML and DL algorithms working from adequate metadata extraction with no effort of decryption. Taking into account the ciphering issue, an analysis based on fixed pattern which is usual to detect malicious behavior is not possible. The methods have to rely on metadata such as packet lengths, timestamps, directions and inter-arrival times. They could also be supported by failures in the protocol handshake allowing mainly key sharing methods in the establishment of the encrypted communications.

The network inspection functionality will be first of all highly dependent on the traffic type identification and on the metadata extraction capacities allowing it. Independently of the following updated encrypted traffic classification SOTA, collecting the metadata and extracting appropriate features to feed the efficient learning methods will be highly dependent on the chosen C4IIoT architecture fixing the context of the traffic.

The traffic classification will be applied depending on the contextual architecture of C4IIoT to make an identification of the type of traffic, before using the partners' technology to perform the final malicious behavior detection. This is applied to the mobile applications in the reference [6] of the up-dated SOTA and seems promising while it is not demanding too many processing resources and is adapted to virtualized software-based environments. The detection of a malicious behaviour is ultimately specific to the application and the malicious action such as described in the different references 3,4,6 of this paper

3.2 Overview of research advances since August 2018

Anonymous communication technologies, such as anonymous proxy and anonymous routing are frequently used to ensure the security of network terminal and personal privacy. In the context of the rapid deployment of these technologies, one can find recent papers, and some of them, published after August 2018, include also SOTAs of ciphered traffic classification, either specific to their technology or more generic. In [7], it concerns the anonymous routing technology Tor, in [8] it concerns the anonymous proxy technology ShadowSocks.

Tor (The Onion Router) is the popular anonymity service that improves privacy and security over Internet. Onion routing is a concept of anonymous communication over a computer network where the messages are encapsulated in multiple (currently 3) layers of encryption.

In [7], the specificity of Tor encrypted traffic regarding other encrypted traffic is the lack of attributes of type: source/destination IP address and port numbers to learn and classify the traffic accurately, as these attributes are hidden by the onion routing mechanism. Hence, other attributes, based on the specific fragmentation of the Tor traffic layers (TLS based) are used for classifying the Tor encrypted network traffic.

Still in [7], a detailed SOTA of the papers which describe ML methods used for traffic classification is presented in chronological order from 2012 to 2018. The reference [6] is part of it.

In [8], the method of identification of traffic in Shadowsocks is presented. Shadowsocks is a free and open-source encrypted proxy project, widely used in China to circumvent Internet censorship. Typically, the client software open a proxy on the machine it is running, which internet traffic can then be directed towards, similarly to an SSH tunnel. Unlike an SSH tunnel, Shadowsocks can also proxy UDP traffic. The usual feature-based methods of identification that rely on traffic characteristics (e.g. extracted from protocol handshake in the initial stage of the establishment of encrypted communication) can't be efficient in the case of the Shadowsocks traffic. Features specific to the Shadowsocks architecture are used to identify the traffic (Shadowsocks flow or not). They are divided in 3 categories: the flow context (roughly quantity of flows correlated with the flow in a given time interval), the source-side host behaviour on the flow and on the DNS.

In [9], one can find the most complete and detailed SOTA related to the Machine Learning solutions to classify network traffic including ciphered traffic. One first explains why ML methods are going to replace the methods which are port-based and even the methods DPI-based. The study is focused on the traffic classification at the IP level. It presents the whole picture of the steps needed for network traffic classification.

It underlines the successful key factors: a reliable label assignment for the construction and validation of the ML models, a dynamic feature selection to create adaptive models that use the most suitable features given the context and the objective to achieve, the integration of meta-learning processes for dealing with the imbalance and the dynamism of the Internet network data and the strategies for the online reconfiguration of the ML solutions.

In [6], the work starts from the observation that traditional approaches for traffic classification were designed to work on a dedicated hardware at very high line rates and may not function well in a virtual software-based environment. It devises a novel fingerprinting technique that can be utilized as a software-based solution which enables machine-learning-based classification of ongoing flows. The scheme is simple to implement, requires minimal resources and attains very high accuracy. Moreover, its performance is essentially independent of placement and migration issues, network conditions such as congestion, fragmentation, delay, retransmissions, duplications, and losses and to varying processing capabilities, and thus yields an attractive solution for virtualized software-based environments.

Specifically, for TCP flows, the fingerprinting scheme is based on zero-length packets i.e., packets that contain control bits, but do not contain any payload (e.g., SYN, ACK, etc.).

Results show that the scheme correctly classified about 97% of the flows on the dataset tested, even on encrypted data.

In [10], one presents a solution to analyze the behavior of a user accessing web services through http/2 over TLS. It detects if a user performs an action previously defined over a monitored web service.

3.3 Overview of innovative commercial products released since August 2018

According to the Gartner source, by 2019, 80% of all traffic is encrypted and 70% of network attacks takes advantage of ciphered traffic. Therefore the big actors of traffic analysis have to enforce their commercial products to take into account this big issue.

Blake Anderson Advanced security research group of CISCO and author of the reference [33] in the proposal, presents in 2019 the “Network visibility and security analytics: Cisco Stealthwatch release 6.9.2¹”.

“Encrypted Traffic Analytics” is a new technology of CISCO² that takes advantage of a dedicated ASIC architecture to extract the data elements without slowing the data network, The collection of metadata in use consists of the sequences of packets lengths and the times between them. This enhanced telemetry is passed to Stealthwatch which applies multiple analytics techniques to detect malware with fidelity

¹<https://www.cisco.com/c/en/us/products/security/stealthwatch/index.html>

²<https://www.cisco.com/c/dam/en/us/solutions/collateral/enterprise-networks/enterprise-network-security/at-a-glance-c45-740079.pdf>

4 Vulnerability analysis and mitigation

4.1 Proposal's stated offering beyond August 2018 SotA

In terms of criticality assessment, current methods use simple dynamic analysis. For example, they find input bytes in the program counter value to assess the impact of a bug [11]. Also, formal proofs are currently done at source level, or sometimes at assembly level with high workload from the user. In particular, current binary-level program analysers suffer from many hardcoded limitations (trade exhaustiveness for efficiency or precision). In terms of automatic patch generation, current work only looks for bugs [12]. Finally, current solutions cannot offer continuous verification, which will be provided in C4IIoT Figure 2.

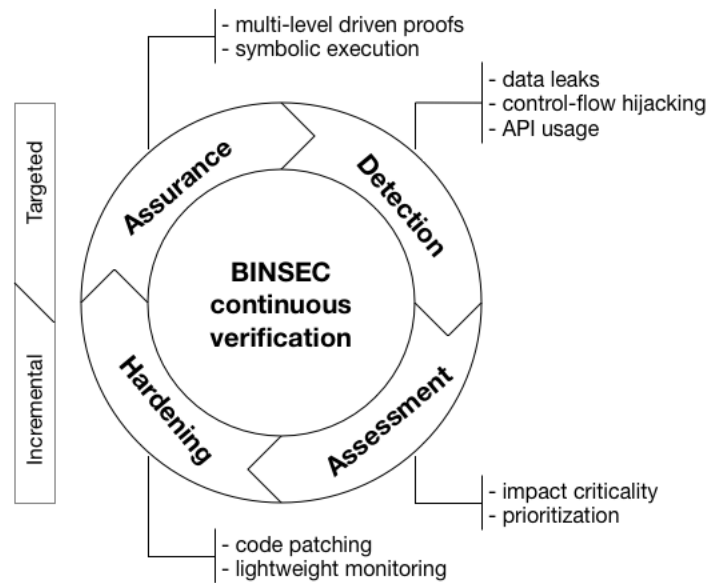


Figure 2. C4IIoT Continuous Verification

Automated assurance, continuous monitoring, smart evidence aggregation and evaluation have recently been hot topics for both industry and academia. Koschorreck [13] focused on automated audit and security controls with particular attention to models and standard protocols to represent and exchange audit results and operations. Several works also exist on security and privacy assessment based on compliance ([14], [15], [16]) and audit ([17], [18], [19], [20]).

Going beyond the state of the art, C4IIoT will develop a semantic framework to have more precise reasoning (taint, constraint solving) in line with preliminary work [21] on the topic. It will also offer automated binary-level proof, and automatic patch generation, going beyond the state of the art to generate "likely-patches" for the easiest cases, keeping human involved only for the most complicated cases. C4IIoT is considered a complex system where continuous security monitoring and assessment is a challenging research problem with no definitive solution as of yet. There are several solutions that partially address this problem, but none of them provide a holistic and disruptive security-enabling solution for prevention & protection against attacks targeting modern IoT components, complex Industrial IoT infrastructures and emerging technologies. In this context, C4IIoT will provide such a solution driven by STS's security assurance and dynamic certification platform. This solution

aims to cover the variability of the different scenarios maintaining the same efficiency in monitoring and assessment of security.

4.2 Overview of research advances since August 2018

In terms of more recent advances, Anisetti [22] presents a solution for IoT security monitoring and assessment considering different attacks surfaces and attacker capabilities and exploiting well-known attacks after a device discovery phase. In this context, Moon Cloud is also presented, for ICT security governance in IoT and Cloud environments [23]. Farinier et al. [21] focused on generating models of quantified first-order formulas over built-in theories that allows to reuse all the existing methods for the software verification and bug finding context. More specifically, their proposed approach alters the quantifier formula into a quantifier-free formula and assures that any model of the latter contains a model of the former. Farooq [24] introduced an analytical model to study device-to-device propagation of malware in wireless IoT networks. Khodjaeva et al. [25] propose several solutions that enable RFID technology to be more secure and be utilized by existing IoT devices without having concerns regarding security, privacy and trust. Nadir et al. [26] formulated a modular approach to implement an auditing framework for the security of IoT devices including hardware, firmware and communicating vulnerabilities. Gemini and Thampi [27] presents a solution for edge devices in IoT where a multi-attacker multi- target graphical model for risk assessment is introduced. Rohit et al. [28] present a testbed for mitigation and vulnerability assessment of industrial components that are being used in real critical systems and mitigation techniques that they have deploy to defend against the discovered attacks. Lastly, Speicher et al. [29] proposes an approach of conducting what-if analyses to reason about mitigation.

4.3 Overview of innovative commercial products released since August 2018

No new commercial products could be identified that were released since August 2018.

5 Decentralized selective access management

5.1 Proposal's stated offering beyond August 2018 SotA

In C4IIoT, we will develop a decentralized solution using Blockchain technology that enforces privacy-aware policies and data access restrictions. Since decentralized solutions built on Blockchain technologies do not scale yet to the amount of data aggregated by IIoT devices and in order to support the “right to be forgotten”, we will develop a hybrid solution. In this solution, IIoT aggregated data and analytics results will be stored on a restricted secure storage on the cloud, after encrypting the data. Records containing links to the restricted storage will be recorded on the Blockchain by the decentralized access management solution. Hash of the data will be stored on the Blockchain as well, to allow verifying data integrity on the cloud. Privacy-aware access policies will be supported using the encryption and decryption mechanism. When a privacy policy or a business decision is changed, the decentralized access management solution will revoke the access to the restricted storage to data federation partners as required. In addition to enforcement of policies, this solution will leverage the records documented on Blockchain to provide forensics and to prove compliance for audit purposes.

5.2 Overview of research advances since August 2018

Several papers were published since August 2018 introducing research advances in the domain of decentralized access control and privacy-aware policies in the world of IoT.

A partial decentralized access control solution for IoT devices is presented in [30]. The authors propose an architecture enabling IoT devices' owners to expose resources to be used by other entities. The solution uses attribute-based access control (ABAC) where resource owners define attribute-based policies to allow or deny access to resources. The Tangle, which is a distributed ledger claimed to overcome some scalability limitations of the Blockchain, is used to store access policies by the resource owners. When a requester wishes to access a resource, his attributes are verified, and the relevant policy is being read from the Tangle and evaluated by a central authority called the “policy decision point”. The decision is sent to the resource owner, who in turn records it in the Tangle to be used to grant access to the resource. The proposed mechanism takes advantage of the distributed ledger technology to audit access policies and access requests and decisions, allowing verifiable flow. Another advantage is the use made in ABAC, allowing fine-grained privacy-aware policies to be defined by the resource owners themselves. However, the mechanism requires a trusted authority to be involved in every access request, evaluating the access policy. Furthermore, it requires approaching the resource owner as well in every access request. In C4IIoT we aim to introduce a solution that allows higher level of decentralization and reduced on-line involvement of central authority.

A privacy-preserving mechanism for sharing and distributing data for IoT, with decentralized elements, is presented in [31]. This paper focuses on the domain of healthcare IoT and uses hybrid approach combining both Blockchain and Cloud Storage used to transfer and store personal data originated from IoT devices. The authors present a solution where patients with wearable IoT devices can share personal data with other parties, e.g. healthcare providers, by sending it to a Blockchain platform that uses smart contracts. The Blockchain platform is not used for storing the data but rather for analyzing it. The smart contracts execute policies in which when a certain condition regarding the content of the data is met, e.g. abnormal

measures appear, an alert will be sent to an agreed-upon data consumer, and the data will be encrypted and stored in the cloud. A complex encryption mechanism that involves both the Blockchain and the cloud is in the core of the solution. The public key of the specific data consumer is used in a two-phase encryption process, thus assuring only the authorized consumer can decrypt and view the data. This is also the main drawback of the proposed solution, as every such a session of data sharing by the patient is aimed at a specific data consumer. In C4IIoT we aim to support the distribution of data to multiple entities, according to privacy-aware policies, and without the edge IoT device having to specify or even be aware of any specific entity in the access policy.

5.3 Overview of innovative commercial products released since August 2018

IBM has a mature Data Policy and Consent Management (DPCM) solution [32]. DPCM is a data governance decision engine that models and links complex privacy-aware policies and consents to actual data items. It can be used by companies to define data policies and collect consents from data subjects thus allowing GDPR compliance. Companies may use DPCM when inserting enforcement points throughout the system to check legitimacy of data accesses. DPCM is a solution constantly under development, with August 2019's version 3.10 being its latest version. This version includes extended reporting functionality, providing companies with information and business analytics regarding the consents given by data subjects to the various purposes and data categories defined, to be used to track and prove compliance. It also presents optimizations allowing improved system scalability. IBM's DPCM is a state-of-the-art in its domain, but it is a fully centralized solution, requiring data subjects to trust the entity managing their data to use DPCM properly and follow by DPCM's data access decisions. In C4IIoT, IBM will provide a privacy-aware data access management solution with decentralization as a core element.

CPChain (Cyber Physical Chain) is a company that develops "a new distributed infrastructure for next generation IoT" [33]. The company's website mentions a few planned releases of relevant products during 2018-2020. In its white paper published on June 2019 [34], it describes its solution for a decentralized data platform for IoT devices that allows storing and sharing of data. The system consists of several layers and presents a hybrid approach to store and track data. Raw data is stored on a cloud storage platform constituting part of the data layer, after being encrypted by the user owning the data before uploading it. A control layer is making usage of the Blockchain technology to store hash records of the data stored in the cloud, thus providing a way to track and verify the integrity of data stored in the cloud while avoiding the storage-costly practice of storing the actual data on the Blockchain. The data access control mechanism is based on encryption technology. A combination of symmetric encryption to encrypt the data itself, and an asymmetric encryption to encrypt the symmetric key and transmit it to a certain data consumer, is suggested. However, the problem of allowing "one to many" authorization to access data in contrast to "one to one" authorization is presented as an open issue yet to be addressed. Overall, the system consists of a few elements that IBM plans to introduce to its C4IIoT contribution as well, but the privacy-aware access control mechanism IBM aims to provide, while maintaining decentralization, is beyond what is presented in the CPChain solution.

6 Dynamic verification and re-configuration of self-adaptive systems

6.1 Proposal's stated offering beyond August 2018 SotA

Self-Adaptive Systems (SAS) [35] [36] are systems able to autonomically [37] detect persistent changes in their operational execution context and adapt their behavior to continue meeting their requirements in the new context. With respect to cybersecurity, context changes can include an alert of a new security threat that is spreading through the network or detection of a highly-probable intrusion occurrence. Such adaptive ability requires access at run-time to explicit models of (a) the variability of the operational context (*context model*) (b) functional and non-functional, hard and soft requirements (*requirement model*) (c) a variety of reusable realization assets (*e.g.*, architecture models, code files) that can be automatically configured and composed in many different ways to satisfactorily implement these requirements in most to all possible contexts (*asset model*) and (d) the complex constraints relating contexts, requirements, configurations and assets (*constraint model*). It also requires an inference engine to automatically reason about these constraints to (a) verify after each context change whether the current configuration still satisfactorily satisfies³ its requirements in the new context and when it does not (b) find a new configuration that does. It must also include an implementation platform mechanism (*e.g.*, hot code push) allowing the executable code of the newfound configuration to replace the old one without interrupting the execution of the SAS.

Software-intensive systems Product Lines Engineering (SPLE) [38] is a discipline to engineer large families of systems that share a set of core features and configurable components while also differing in terms of other features and components. In SPLE, a *feature* is defined as a user-visible set of requirements that is coherent from marketing, regulatory and technical perspectives. A traditional SPL shares with a SAS the need for a requirement model, an asset model and a constraint model. The requirement model generally takes the form of a *feature model* consisting of a feature decomposition tree plus constraints relating features across tree branches. SPLE commercial tools (*e.g.*, pure:variants [39] and Gears [40]) and academic open-source tools (*e.g.*, FeatureIDE [41]) include a configuration user interface and a product derivation that automatically assembles the relevant assets from the asset model that together allow implementing the features chosen by the configuration. The academic tools additionally support automated verification of constraints inside and across some models of an SPL. A traditional SPL is *static* in the sense that configuration and product derivation occurs once at development time. The chosen configuration then remains unchanged during the operations of the systems.

Dynamic SPLE (DSPLE) [42] is a recent idea that reuses the rich concepts and tools of SPLE for *SAS Engineering (SASE)* by extending them with a context model, constraints between the context model and the feature, asset and configuration models, as well as hot code push mechanisms. Given the central role of models in SPLE, SASE and DSPLE, they can all be considered *Model-Driven Engineering (MDE)* [43] approaches. Given their need to exploit models at run-time, SASE and DSPLE both belong to the *Model-at-Runtime (M@RT)* [44] sub-field of MDE.

³ <https://en.wikipedia.org/wiki/Satisficing>.

In August 2018, there was no commercial DSPLE tool available. The most comprehensive academic DSPLE tool was VariaMos [45] developed by C4IIoT partner UP1PS. Its main limitations to be overcome during C4IIoT were the following. First, it lacked a model for source and executable code assets needed for automated derivation of running systems. Second, it was architected as a monolithic, stand-alone Java desktop application, not encapsulating in separated components the concerns of persistent models management, automated reasoning and human user-interface while also not providing an *Application Programmable Interface (API)* to access its service from external software. This second limitation makes it hard to extend and integrate with other software in the C4IIoT framework. Third, its built-in requirement modeling language REFAS [46] is general purpose, not offering an out-of-the-box an ontology of cybersecurity hard and soft requirements concepts.

6.2 Overview of research advances since August 2018

Since August 2018, five relevant research papers have advanced the SotA in DSPLE in a way that could be relevant for C4IIoT. Mauro et al [47] presented HyVarRec a DSPLE tool that provides a set of services similar to those provided by VariaMos. The main novelty of HyVarRec as compared to VariaMos is that its context and configuration models keep a history of the changes that both underwent during the operations of the DSPL. However, it does not include any asset model and thus does not support product derivation from a chosen configuration. Another limitation is that it relies on a single inference engine, the Z3 *Satisfiability Modulo Theories (SMT)* solver [48]. In contrast, the next version of VariaMos, intends to adopt as intermediate language between the modelling components and the reasoning components, the MiniZinc [49] standard language accepted as input by numerous solvers. This approach should allow VariaMos to choose the best solver for each different DSPLE reasoning task.

Krieter et al [50] proposed to use an SPL to modularize a C code application into minimal functional features which realization code and needed data are small enough to fit inside the limited memory of encryption protected enclaves managed by the hardware-based security extended instruction set of Intel's SGX processor. Though evaluated only for a single kind of realization asset and a single kind execution platform, this is a very original application of SPLE, startlingly different from its original functional and non-functional variability management purpose. It points to an interesting direction to explore to leverage in synergy in C4IIoT, on the one hand, hardware-protected secure and trusted execution tools such as those surveyed in sections 2 and 7, and on the other, DSPL modelling with tools such as VariaMos.

Concerning the latter, Correa et al [51] proposed a versatile, hybrid asset model for DSPLE called *Fragment-Oriented Programming (FragOP)*. Implementing FragOP asset modelling is on the roadmap of future versions of VariaMos. Historically, SPL asset modelling has been done following one of two broad categories of approaches: *annotative* (also called *subtractive*) or *compositional* (also called *additive*). An annotative approach adds annotations, such as pre-processor statements or comments, to the source code, that tag each code block as the asset realizing a certain feature in the feature model. The product derivation then consists of compiling only the blocks annotated with the features selected by the configuration. This requires an annotation assistant tool to verify that the subtraction of unselected features results in code that compiles. A compositional approach associates features not with arbitrary blocks but only with those encapsulated into reusable components or modules. Derivation then consists of assembling the components realizing the features selected by the configuration. Annotative SPLE allows arbitrarily fine-grained assets and features and is initially simpler to implement. However, it leads to more tightly coupled, harder to understand and hence less

maintainable asset bases than compositional SPLE. FragOP aims to get the best of both worlds by a hybrid approach where features can be associated with two types of assets: coarser-grained components and finer-grained *fragments*. The later are blocks that can extend, substitute or delete special blocks inside components annotated as *fragmentation points*. This is inspired by *Aspect-Oriented Programming (AOP)* but it is language-independent and with a more principled, intentionally constrained expressiveness to prohibit the re-introduction of non-local dependencies between arbitrary blocks.

Weckesser et al [52] present a DSPLE framework (let us call it WEADSPLEF) that uniformly represents the requirements, context and constraints models of a SAS with the SPL concepts of feature trees with feature attributes augmented with Boolean, integer and real valued cross-tree constraints. They however distinguish constraints relating context feature attributes with soft requirements feature attributes, calling them performance-influence models. In WEADSPLEF, these contextual performance constraints are to be machine learned from simulations and/or real-world datasets using a stepwise linear regression algorithm with iterative heuristic forward feature selection [53], while all the other types of cross-tree constraints are to be manually specified. In WEADSPLEF, the task of finding a system configuration that both satisfies all hard requirements and optimizes the soft ones in the current context configuration is viewed as a task of solving a *Mixed Integer Linear Programming (MILP)* problem using IBM's CPLEX solver⁴. Just like HyVarRec, WEADSPLEF does not include a generic asset model. They instead focus on experiments measuring the efficiency of finding a valid and optimal feature configuration given a context configuration.

The current REFAS DSPL modelling language supported by VariaMos can express much of the same models than those of WEADSPLEF. However, in terms of vocabulary, it is less uniform and parsimonious since its concepts are inspired not only from SPLE but also from *Goal-Oriented Requirement Engineering (GORE)* [54]. While it can in principle perform the same kind of optimal feature configuration search than WEADSPLEF, the scalability of its current implementation for this task has not yet being evaluated. Weckesser et al's idea of machine learning constraints relating context and soft requirement variables is an interesting direction to explore to integrate in C4IIoT DSPLE tools such as VariaMos with machine learning at edge tools reviewed in section 8.

Lara et al [55] present a survey of adaptive security solutions based on IBM's MAPE-K reference architecture for autonomic computing which decomposes a SAS into five top-level components:

- The *Monitor M* which senses the operational context of the SAS to detect changes in it;
- The *Analyser A* which evaluates whether the change detected by *M* requires the SAS to change its configuration to still satisfy its fixed requirements or to change its context-dependent requirements;
- The *Planner P* which searches for a new configuration that does satisfy the fixed requirements and the new context-dependent requirements after *A* evaluated that the current configuration does not;
- The *Executor E* which changes the SAS implementation at run-time to conform to the new configuration found by the *P*;
- The *Knowledge base K* shared by *M, A, P* and *E*.

A DSPLE MAPE-K approach further decomposes *K* into a SAS feature tree, a context feature-tree, cross-tree constraints and an implementation asset model. Some of these

⁴ <https://www.ibm.com/analytics/cplex-optimizer>.

approaches surveyed in [55] could serve as inspiration to extend REFAS with built-in cybersecurity concepts.

6.3 Overview of innovative commercial products released since August 2018

During the last year, no commercial DSPLE tool has appeared on the market. The two leading commercial SPLE tools, pure::variants and Gears still only support engineering *static* SPL.

7 Secure-by-design IIoT device fabrication

7.1 Proposal's stated offering beyond August 2018 SotA

Based on the definition of the future system to be deployed, fabrication of novel IIoT devices will be considered (secure-by-design device fabrication), relying on LP-WAN communication technologies, such as NB-IoT or LoRA, mainly in collaboration between VIP and CRF. The design of IIoT devices will ensure supporting advanced hardware-enabled security features, from IFAG, but will also account for memory/processing requirements and incurred battery consumption costs (in case of battery-operated nodes). The joint pool of fabricated devices and existing CRF IIoT devices will be used, tested, and evaluated both in simulated and in real-world environments. A special attention will be given to testing and evaluating the security-enabling capabilities of the devices. The comprehensive simulated environment will model both the C4IIOT's edge nodes layer and the field gateway layer. This will be accomplished within a virtualized environment at a dedicated server (visible through a dedicated IP address) and residing in proximity of suitable mobile core network elements. The results of testing and evaluation will be used to fine-tune the developed solutions.

Current commercial IoT modules provide limited processing/memory/battery capacity resources to support a number of standard security algorithms based on encryption methods. In addition, running hardware-based security tests and integrity checks could provide further strain on available resources. Even if used for standard unencrypted data transmission, most of the commercially available platforms are not optimized for specific IIoT applications and would typically deplete the battery in a short time period. Besides, their form factors are very often not a good fit for a desired application. Examples of such platforms range from Arduino Uno, Intel Edison or Raspberry Pi, equipped with communication modules for specific LP-WAN wireless technology. For example, the state-of-the-art commercially available IoT development solution for NB-IoT technology is Sodaq NB-IoT shield for Arduino, which is based on Ublox Sara 211 NB-IoT modules. However, some drawbacks of this module are identified by the UNSPMF group while working with this module in a live operator NB-IoT network, e.g. it prevents using this IoT platform beyond purely educational and training purposes.

In C4IIOT, we plan to innovate beyond the state-of-the-art in the “secure-by-design” field. We will make a systematic attempt to provide requirements on IIoT module resources to enable different levels of security matched to different IIoT application requirements. The project will use the strong experience of the UNSPMF group in past years in the design and fabrication of IIoT modules relying on LPWAN communication technologies, such as NB-IoT or LoRA, for specific use case deployments and demonstration purposes in collaboration with Serbian and Danish mobile operators. C4IIOT will design and fabricate IIoT modules capable of supporting advanced hardware-enabled security features, but will also carefully investigate their memory/processing requirements and incurred battery consumption costs. For the latter, we will also consider integration of energy harvesting modules on IoT platforms in order to reduce dependence on battery resources. Overall, different IoT device designs will be provided that would allow testing and evaluation of progressively stronger hardware-enabled security routines developed within C4IIOT project.

7.2 Overview of research advances since August 2018

To reach a higher security level of IIoT does not only depend on the industrial players or users. The lawmakers and overall regulations would help to set up a transparent and fair framework and guide the improvements. UK government “Department for Digital, Culture, Media & Sport” has collected best practice examples from the consumer IoT devices, manufacturer, based on these the handbook “Secure by Design” [56].

Various surveys show clear evidences that there is still a lot of room to improve IoT devices security. The article “A Survey on Hardware-based Security Mechanisms for Internet of Things” [57] has a review of security challenges of emerging IoT networks, listing various attacks and countermeasures. Regarding threat analysis “Comprehensive Review on the Issues Related to the Data Security of Internet of Things (IoT) Devices” [58] provides comprehensively examples faced by IoT devices. Possible solutions are suggested to solve the issues. Not only hardware is considered for security improvement, also software patches are often necessary, and reliable secure firmware update mechanisms become essential. For that the paper “Hardware/Software Security Patches for Internet of Trillions of Things” [59] proposes a framework of adding integrated hardware and software patches as a security monitoring and protection layer to the existing devices (things). Also, some new approaches about how to tackle information security in IoT are of interest to be studied. “Internet of Things: information security challenges and solutions” [60] provides some considerations and approaches.

Correlated to the various security weaknesses we can observe increasing interest of investment and design improvement. Continuous studying and publications are visible. Nevertheless, we see diversified approaches to improve the system. In overall, the hardware-based security is considered as one of the most important mechanism to counterfeit security weakness and attacks. Representative paper is “Design and Verification Methodology for Secure and Distributed Cyber-Physical Systems” [61]. It describes a new design and verification methodology for secure and distributed microcontroller-based devices. In addition “Next Generation Resilient Cyber-Physical Systems” [62] aims to analyse the future challenges of such systems and based on these derives the requirements.

In conclusion from the observation, we see a wide range of actors from government, research institutes, and companies investing continuously in IoT security. This gives the evidence that the demand for security is very high on one side. On the other side it would require huge efforts and resources to analyse the weaknesses and create concepts against these. The next section summarizes some concrete instances and solutions in market.

7.3 Overview of innovative commercial products released since August 2018

In relationship to the study and scientific researches, many commercial products are introduced to solve concrete issues and improve dedicated weaknesses. The hardware based security is on the main stream.

As one of the leading security providers, MULTOS introduced the branded Chip to Cloud Security [63]. It aims to cloud services and is based on the principle of device trust anchor. Device Authority’s KeyScaler and MULTOS integrated solution delivers a Secure by Design approach focusing on key provisioning and key insertion for initial device trust, device provisioning and operationalizing trust and security operations into the Enterprise. A special chain in the secure key management procedure: MULTOS uses initial key insertion being

offered through the MULTOS manufacturer's process and partners. Typically, Infineon Technologies has Common Criteria certified secure manufacture environment and is thus suitable for such key ceremony.

ARM based architectures are gaining more and more market attention. This is not only visible in mobile handset but also in IoT devices. "Trustonic IoT Developer Kit – Secure By Design" [64] MCU features Arm TrustZone and Trustonic's Kinibi-M Trusted Execution Environment (TEE) and claims to be the first of such utilization. The cooperation of Microchip and Trustonic on advanced hardware-based SAM L11 microcontroller family targets IoT developers and embedded systems developers. They can now build the most secure solutions with the easy to use toolkit. With the growing need to robustly secure IoT devices, services and infrastructure, Trustonic and Microchip solutions combine to enable multiple secure use cases.

Particularly to protect edge devices, Adlink Technology and Entrust Datacard partnered to work on the creation of an Industrial IoT (IIoT) security model, enabling secure communication for data streams throughout the entire IoT value chain - from manufacturing and applications, to endpoints and edge devices. They consider security as one of the most significant barriers to the adoption of the Internet-of-Things (IoT).

We can see a similar statement also from Microchip [65]. Microchip security products make "trust" easy to embed in any system. Flexibility, advanced features, innovative cost-effective architectures and ultra-secure hardware defense mechanisms make Microchip hardware-based security devices an ideal way to add trust by design at scale in various use cases (e.g. IoT cloud authentication, automotive security, counterfeit protection etc.).

An important consideration that some say differentiates IIoT security from traditional IoT security concerns is the life cycle management (LCM). Secure Thingz's Povey said that LCM has an impact on when software updates or configuration changes are deployed to IIoT devices. In IIoT environments, the connected devices, sensors, and control systems will typically not, or should not, be connected to the open internet. But security needs vary in an IIoT network depending on the endpoints in the system because it may comprise both an offline internal network of non-IP-based smart controllers and some type of protection or isolation from the external internet, and there will also be wireless devices and sensors that may or may not be IP-based. All of the endpoint devices need to be managed and controlled in an industrial system as part of the LCM function [66].

As conclusion, the integration of hardware based security, and a complex and tightly managed life cycle (from device manufacturing and root-of-trust generation to long-term support) is a very important approach to improve "IoT security-by-design". It would require a deep cooperation in a complex and big value chain: it contains hardware IC, software IP like ARM, system integration, communication modules, key generation, distribution and management, etc. It is almost only manageable beyond a single entity in the current age, and trust between different manufacturers at different layers of the chain is needed, what highlights the importance of trustworthy players and consortiums to build trust, such as the one for this project, or to create trustless-alternatives, based on blockchain approaches. It reflects also the observation that counterfeit and IoT threats are a cross-functional activity.

8 Machine learning at the edge

8.1 Proposal's stated offering beyond August 2018 SotA

The emerging concept of Edge Intelligence (EI) [67], refers to edge computing with machine learning and advanced networking capabilities. EI relates to the trend that several information technology and operational technology industries are moving closer towards the edge of the network so that aspects such as real-time networks, security capabilities to ensure cybersecurity, self-learning solutions and personalized/customized connectivity can be addressed [68]. More specifically, there has been a very recent effort in moving not only inference of a ML algorithm, but also its training, closer to the edge. Current and now-emerging solutions consider the usage of “containerized” ML software modules that perform standard EI services and are deployed at the edge devices. From the ML algorithmic perspective, they primarily utilize the concept related with federated learning [69], where each edge device downloads a global model from the cloud, performs model update based on its local data (hence not sending the raw data to the cloud), and uploads its local model refinements to the cloud. In other words, from the algorithmic point of view, *data partitioning* is performed across all edge devices' data to update the *global model* stored in the cloud.

In C4IIoT, we take a step towards simultaneous data partitioning and model partitioning of ML algorithms [70]. For model partitioning, we bring a ML model (e.g., a deep neural network) closer to the edge. More precisely, we will examine multiple ways to partition the ML model across the three layers of the C4IIoT architecture – the edge nodes, the field gateway, and the cloud. This corresponds for example to update the localized model parts attached to a group of edge devices at the physically close field gateway and not the cloud. Such modelling shift potentially significantly reduces time delays and increases algorithm robustness, making the ML algorithms more amenable for anomaly detection tasks. Specifically, we will exploit several ML models and algorithms approached, including deep neural networks, structured (non)convex optimization models, and belief propagation-type algorithms [71].

Taking into account the envisioned architecture of the C4IIoT framework, a distributed machine learning mechanism for anomaly detection (AD) could be achieved by a hierarchical collections of machine learning models with various degrees of complexity and data/model partitioning: (1) a standalone, lightweight AD model deployed at edge nodes that could be updated from higher-levels AD models, (2) an AD model with data and model partitioning across edge nodes receiving inputs from edge nodes AD models that is deployed at field gateways, and (3) cloud-deployed AD model with data and model partitioning across the cloud, edge nodes and field gateways receiving inputs from and updating AD models at the lower levels. It should be emphasized that the cloud-level AD model can be partitioned only across the cloud in the case that edge nodes cannot store a complex AD model (which is case with data partitioned complex machine learning models) or when partial updates of model parameters with local data degrade AD performance (in case of model partitioning).

A supervised decentralized AD model can be realized only if edge nodes contain or produce labelled (annotated) data and network flows. In this case, training data for higher-level AD models can be generated from lower-level AD models, i.e. edge node AD classifiers generate training data for AD models deployed at field gateways and in the cloud. Otherwise, unsupervised decentralized AD models should be employed. Unsupervised decentralized AD could be achieved by one-class classifiers which learn normal network behaviour at different

IIoT layers. This direction could be very promising having in mind recent advances in deep learning based one-class classification (e.g. one-class generative adversarial networks by Schlegl et al. proposed in 2017, deep support vector data description proposed by Ruff et al. in 2018, one-class convolutional neural networks proposed by Perera and Patel in 2018). However, the proposed deep one-class classification models should be adapted for federated learning settings, which could be another C4IIoT innovation.

8.2 Overview of research advances since August 2018

Information technology (IT) and operational technology (OT) industries are moving closer towards the edge of the network so that aspects such as real-time networks, security capabilities to ensure cybersecurity, self-learning solutions and personalized/customized connectivity can be addressed [67]. Mission-critical applications such as factory automation, self-driving cars, facial and image recognition require not only ultra-low latency but also high reliability and fast, on-the-fly decision-making. Any centralized architectures are not able to provide the new performance requirements mostly due to congestion, high latency, low bandwidth and even connection availability. Furthermore, fast decision-making on the edge needs advanced computing capabilities right on the spot, which can be provided only by onboard computers or interconnected edge-computing local nodes working together and this makes it very expensive. Machine Learning on the edge alleviates the above issues and provides other benefits. From the cybersecurity perspective, it is a tremendous opportunity to apply AI and machine learning into “edge computing & IoT” to better analyse different cyber behaviours, identify potential threats and vulnerabilities for repairing or patching, and detect malicious attacks [72].

A Harvard Business Review [73] revealed that two areas where AI can have the greatest impact are the retail and advanced manufacturing. AI can create \$1.4 trillion to \$2.6 trillion of value in marketing and sales across the world's businesses, and \$1.2 trillion to \$2 trillion in supply-chain management and manufacturing. In manufacturing, the greatest value from AI can be created by using it for predictive maintenance (about \$0.5 trillion to \$0.7 trillion across the world's businesses). AI's ability to process massive amounts of data, including audio and video, means it can quickly identify anomalies to prevent breakdowns, whether that be an odd sound in an aircraft engine or a malfunction on an assembly line detected by a sensor. While, Forbes indicates that by 2021, 20% of leading manufacturers will rely on embedded intelligence, using AI, IoT, and blockchain applications to automate processes and increase execution times by up to 25%.

With respect to research efforts, several papers were published since August 2018 introducing advances and further research ideas in the domain of machine learning at the edge. In a study published early September 2018 [74], tests were being conducted comparing a number of ubiquitous machine learning algorithms on IoT edge devices. Specifically, the study compared the performances of all three algorithms namely Multi-layer perception, Random Forest and SVM algorithms. The Random Forest algorithm proved to be slightly faster in speed and widely better in accuracy. However, looking at the research from a wider perspective, all of the algorithms' accuracy exceeded 80%, the time required to run them for inference was below one millisecond and they all had moderately low energy consumption. Hence, the conducted research proves that running the state-of-the-art machine learning algorithms is feasible to be run on edge IoT devices for all purposes. As a recommendation, the idea of implementing more complex and taxing algorithms such as Deep Learning, using platforms like TensorFlow, on these small devices would be the next step in revealing their

power in more detail. Work on deployment of pruned deep learning models in recent research looks promising. [74] [75]

Equally important, towards the end of 2018 a set of new design principles to the wireless communication community were introduced towards an upcoming era of edge intelligence [76]. The introduced learning-driven communication techniques can break the communication latency bottleneck and lead to fast edge learning, which are illustrated in three key topics: computation-oriented multiple access for ultra-fast data aggregation, importance-aware resource allocation for agile intelligence acquisition and learning-driven signal encoding for high-speed data-feature transmission. Moreover, in the same research it is highlighted that cloud learning and edge learning can complement each other with their own strengths. The federation between them allows the training of more comprehensive AI models that consist of different levels of intelligence. For example, in the industrial control application, an edge server can be responsible to the training of low-level intelligence such as anomaly detection, for tactile response to the environment dynamics. On the other hand, a cloud server can concentrate on crystallizing the higher-level intelligence, such as the regulating physical rules behind the observations, for a better prediction of the ambient environment. More importantly, the collaboration between the cloud and edge learning can lead to mutual performance enhancement. Particularly, the performance of the low-level edge AI can be fed back to the cloud as a learning input for continuously refining the high-level cloud AI. In return, the more accurate cloud AI can better guide the model training at the edge. Nevertheless, how to develop an efficient cooperation framework with minimum information exchange between the edge server and cloud server is the core challenge that need to be addressed. [76]

Recent research efforts in the field have been focused on how to efficiently utilize the limited computation resources at the edge for the optimal learning performance of machine and deep learning models. Wang et al. (2019) proposed a control algorithm that determines the best trade-off between local update and global parameter aggregation in data partitioned federated learning models trained using gradient-descent algorithms [77]. In a research published in May 2019” [78] it is stated that most of deep learning-based AI models are highly resource-intensive, which means that powerful computing capability supported by abundant hardware resources (e.g., GPU, FPGA, TPU) is an important boost the performance of these AI models. Therefore, there are many studies to exploit model compression techniques (e.g., weight pruning) to resize the AI models, making them more resource-friendly for edge deployment. Along with a different line, resource aware edge AI model designs have been considered; Instead of utilizing the existing resource-intensive AI models, the AutoML idea [79] and the Neural Architecture Search (NAS) techniques [80] are leveraged to devise resource-efficient edge AI models tailored to the hardware resource constraints of the underlying edge devices and servers. Methods such as reinforcement learning, genetic algorithm, and Bayesian optimization can be adopted to efficiently search over the AI model design parameter space (i.e., AI model components and their connections) by taking into account the impact of hardware resource (e.g., CPU, memory) constraints on the performance metrics such as execution latency and energy overhead.

Important research advances were also made regarding deep neural network architectures for resource-constrained devices: Zhang et al. proposed an extremely efficient convolutional neural network for mobile devices and Nikouei et al. introduced a lightweight convolutional neural network that can run on edge devices [81]. The literature review by Hussain et al. (2019) [82] summarizes security challenges and threats models for IoT networks and lists current applications of machine learning (ML) and deep learning (DL) techniques in the field of IoT security. More specifically, ML and DL techniques can be employed for authentication

and access control in IoT networks, anomaly and intrusion detection, malware analysis and DDoS attacks detection and mitigation. The review also indicates the current limitations of ML/DL models (e.g. scalability issues and resource limitations) and future research challenges. The following research advances in ML and DL techniques for IoT security have been made since August 2018 (including also papers not referenced and covered in the literature review made by Hussain et. al):

- Nguyen et al. (2019) proposed DIoT – a distributed self-learning system for detecting compromised IoT devices. In DIoT, a recurrent neural network (RNN) is trained for each device type present in the IoT network to learn a normal communication profile. A federated (distributed) learning scheme is employed to learn device-type specific RNNs [83].
- Ferdowsi and Saad (2019) proposed a distributed privacy preserving IoT intrusion detection security system based on federated generative adversarial networks. In the proposed decentralized architecture, every IoT device monitors its own data as well as neighbor IoT devices to detect internal and external attacks [84].
- Meidan et al. (2018) proposed N-BaIoT – a method for detecting IoT botnet attacks based on deep autoencoders. For each device present in a IoT network, a deep autoencoder is trained on features extracted from normal traffic data [85].
- Bezerra et al. (2019) proposed IoTDS – a distributed method for detecting IoT botnet attacks based on light-weight one-class classification models (elliptic envelope, isolation forest, local outlier factor and one-class support vector machines) [86].
- Rathore and Park (2018) created a decentralized attack detection framework for IoT networks based on semi-supervised learning employing extreme learning machines and fuzzy C-means algorithms. The authors demonstrated that their decentralized attack detection approach achieved better performance in terms of detection time and accuracy compared to centralized solutions [87].
- Li et al. (2018) proposed an AI-based two-stage intrusion detection system for software defined IoT networks. The proposed approach encompasses the Bet algorithm with swarm division and differential mutation to select typical features and the random forest classification model to classify network flows [88].
- Doshi et al. (2018) employed various machine learning algorithms (k-nearest neighbor, support vector machines, decision trees and neural networks) to detect DDoS attack traffic in consumer IoT devices [89]. Pajouh et al. (2018) proposed a malware detection approach for IoT based on deep recurrent neural networks [90].

8.3 Overview of innovative commercial products released since August 2018

In August 2018, edge AI emerges in the Gartner Hype Cycle for the first time [91]. According to Gartner's prediction, edge AI is still in the innovation trigger phase, and it will reach a plateau of productivity in the following 5 to 10 years. In the industry, many pilot projects have also been carried out towards edge AI. Specifically, on the edge AI service platform, the traditional cloud providers, such as Google, Amazon and Microsoft, have launched service platforms to bring the intelligence to the edge, through enabling end devices to run ML inferences with pre-trained models locally. Below popular ML/DL tools that have been recently been updated (updates within last 12 months) are listed [92]; [93]

Table 1: Popular ML/DL tools

TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets	Updated release:
---	------------------

researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications. TensorFlow includes an implementation of the Keras API (in the tf.keras module) with TensorFlow-specific enhancements.	06/2019
<u>Microsoft Azure Machine Learning Workbench</u> is an end-to-end data science and analytics solution that helps professional data scientists to prepare data, develop experiments, and deploy models in the cloud. Azure Machine Learning is an existing service that helps data scientists and developers build and train AI models more rapidly, and which streamlines their deployment to an edge appliance or the cloud. The new features (updated: Sept 2018) include tools to automate the process of building a machine learning model, by helping to identify the most efficient algorithms to make a prediction and optimize the performance of the trained model. There're also new hardware-accelerated models for FPGAs, and a Python SDK that makes Azure Machine Learning services accessible from popular IDEs and notebooks.	Updated release: 08/2019
<u>Microsoft Azure Sphere</u> aims to secure connected microcontrollers at both the board level and network level. Azure Sphere has three components. The first is customized microcontroller units (MCUs) for IoT devices, which are authenticated using certificates encoded in on-board chips. The second component is the Azure Sphere OS, which runs on the IoT devices and helps secure and authenticate the hardware, and which is based on a custom version of the Linux kernel. The third is the Azure Sphere Security Service, a cloud-based offering that keeps devices patched with the latest security updates and detects threats to these connected devices for 10 years after their rollout. Azure Sphere-certified development kits are now available.	Updated release: 05/2019
<u>The Microsoft Cognitive Toolkit (CNTK)</u> is an open-source toolkit for commercial-grade distributed deep learning. It describes neural networks as a series of computational steps via a directed graph. CNTK allows the user to easily realize and combine popular model types such as feed forward DNNs, convolutional neural networks (CNNs) and recurrent neural networks (RNNs/LSTMs).	Updated release: 04/2019
<u>Keras</u> is a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano. It was developed with a focus on enabling fast experimentation. Being able to go from idea to result with the least possible delay is key to doing good research.	Updated release 10/2018
<u>PyTorch</u> : Tensors and Dynamic neural networks in Python with strong GPU acceleration. Next to the GPU acceleration and the efficient usages of memory, the main driver behind the popularity of PyTorch is the use of dynamic computational graphs.	Updated release 08/2019
<u>MXNet</u> : Lightweight, Portable, Flexible Distributed/Mobile Deep Learning with Dynamic, Mutation-aware Dataflow Dep Scheduler; for Python, R, Julia, Scala, Go, Javascript and more.	Updated release 06/2019
<u>PaddlePaddle</u> : (PArallel Distributed Deep LEarning) is an easy-to-use, efficient, flexible and scalable deep learning platform, which is originally developed by Baidu scientists and engineers for the purpose of applying deep learning to many products at Baidu.	Updated release : N/A
<u>Ekkono's</u> Edge Machine Learning software is embedded onboard connected devices to make them conscious, self-learning, and predictive. Ekkono provides a highly configurable, small-footprint, platform-agnostic, embedded software library that is built for the purpose to help developers rapidly and easily deploy Edge Machine Learning.	Updated release : N/A
<u>Google Cloud AI</u> provides modern machine learning services, with pre-trained models and a service to generate tailored models.	Updated release : 08/2019
<u>Amazon SageMaker</u> provides every developer and data scientist with the ability to build, train, and deploy machine learning models quickly. Amazon SageMaker is a fully managed service that covers the entire machine learning workflow to label and prepare your data, choose an algorithm, train the model, tune and optimize it for deployment, make predictions, and take action. Your models get to production faster with much less effort and lower cost.	Updated release : 11/2018

<u>The Intel Deep Learning Cloud</u> , or Intel Nervana, is a deep learning framework based on Nervana Systems' Nervana Cloud AI framework, with industry leading performance on GPUs thanks to its custom assembly kernels and optimized algorithms. Intel acquired Nervana Systems in 2016.	Updated release : N/A
<u>SparkPredict</u> from SparkCognition in Austin is a machine learning application, this particular deployment using proprietary cognitive algorithms to supply users with better predictive analytics and faster model building capabilities, notably of use in root cause analysis and design.	Updated release : N/A
<u>Infrd</u> is a cloud-based AI platform with three fundamental capabilities: (i) Computer vision which makes sense of large volumes of images to describe what's in them; (ii) Natural Language processing which makes sense of large volume of text coming from contracts, documents, customer conversations etc.; (iii) Predictive modelling which tries to predict what might happen in the future based on what has happened in the past (and supports the above two capabilities).	Updated release : N/A
<u>Watson Machine Learning Accelerator</u> , a new piece of Watson Machine Learning Accelerator (formerly IBM PowerAI), aims to make deep learning and machine learning more accessible to users. It combines popular open source deep learning frameworks, efficient AI development tools, and accelerated IBM® Power Systems™ servers. The vendor promises that organizations can deploy a fully optimized and supported AI platform that delivers blazing performance, proven dependability and resilience. Watson Machine Learning Accelerator is a complete environment for data science as a service, enabling organizations to bring AI applications into production.	Updated release : 03/2019
<u>IBM Db2 Analytics Accelerator</u> is an appliance which boosts software query performance, helps manage structured and unstructured data, and prepare data for analysis, machine learning, or other advanced tasks.	Updated release : 07/2019
<u>Analance</u> is a robust, scalable end-to-end advanced analytics platform that combines machine learning, artificial intelligence, business intelligence, and data management capabilities in one integrated, self-serve platform. The platform is built to deliver core analytical processing power to ensure data insights are accessible to everyone, performance remains consistent as the system grows, and business objectives are continuously met within a single platform.	Updated release : N/A
<u>SeeDot</u> . In 2019 Microsoft Research India (MRI) created SeeDot – a domain specific language to express ML inference algorithms and a compiler that translates SeeDot programs to fixed-point code that can run on constrained IoT devices. Previously in 2018, the MRI team proposed FastGRNN (fast and tiny sized gated recurrent neural network that can be deployed on resource constrained IoT microcontrollers). Other products released by the MRI team include Bonsai (tree-based classifier for constrained IoT devices) and ProtoNN (prototype-based k-nearest neighbor classifier for constrained IoT devices). All released products are open source and available at https://github.com/microsoft/EdgeML	Updated release : 2019
<u>TensorFlow Federated (TFF)</u> In 2019 Google released TensorFlow Federated (TFF) – an open source framework for machine learning on decentralized data. TFF enables developers to declaratively specify federated computations that could be deployed to diverse runtime environments. The source code of the framework is available at https://github.com/tensorflow/federated . Additional documentation can be found at https://www.tensorflow.org/federated	Updated release : 2019

9 Security-aware offloading decision support

9.1 Proposal's stated offering beyond August 2018 SotA

A year after our proposal submission, the challenges on designing a security-aware dynamic offloading decision mechanism are still relevant and important. Processing, or even producing, the data closer to the **edge** devices is now considered a core enabler of the IoT [94] as well as the 5th generation of mobile communications (5G) [95], since offloading computation **to a nearby device, such as C4IIoT's gateway**, may allow most of the benefits of the cloud without its key communication **latency** disadvantage. The components of IIoT environments, such as Wireless Sensor Networks (WSNs) and Wireless Sensor and Actuator Networks (WSANs), cyber-physical systems, and Web of Things, which enable sensors and actuators to be meshed with services and data on the web, create the need for a middleware in the form of an intelligent gateway or fog that handles resources and communication of the edge nodes and allow close to local processing [96].

At the same time, cyber security is a primary requirement for the adoption of IoT technologies in the industry [97]. Offloading decision support is largely missing in the literature. The few related technologies that exist are still in early stages and typically geared towards manual decision or considering a small subset of the factors that matter in C4IIoT, most notably ignoring the fact that offloading itself increases the attack surface and overheads and needs to be used only when necessary. In C4IIoT our security-aware dynamic offloading decision mechanism will evaluate the trade-off between confidence in anomaly detection and cost of offloading to the cloud (high computational power, high latency), and automatically trigger the latter when appropriate. This allows to run actions at the edge, closer to or on the IoT devices themselves, reduce the attack surface by minimizing communication to the cloud and reduce response times and energy consumption. In this manner, it reduces response times by fully utilising edge computing and reduces security risks and overheads by utilising cloud processing only when needed.

9.2 Overview of research advances since August 2018

In [98], the authors propose a lightweight offloading scheme, where delay-sensitive tasks are given high priority and are executed immediately at the edge while other tasks are offloaded to a remote (Cloud). Thus, the edge only executes delay-sensitive tasks, while the cloud is used for the rest.

Alam et al. [99] propose an offloading mechanism that uses deep reinforcement learning which may introduce response time overheads beyond what is appropriate for IIoT environments. Similarly, the authors of [100] propose to use deep-learning for offloading in order to predict the response times. It predicts the response time of a task based on historical observations regarding CPU, memory, and bandwidth using a Deep Belief Network with 2-3 hidden layers, using pre-define parameters (e.g. bandwidth), which may also introduce noticeable response time overhead due to the decision mechanism, and crucially ignores the cyber security element.

In general, although solutions that are specific to the Industrial IoT environments have begun to emerge, such as [101], they are focusing only on the response times and do not take security or energy consumption into account. An exception is the work by Chen et al. [102], which does take into account energy consumption through an optimisation goal under time

constraint. In this manner it is appropriate for IIoT applications with strict time constraints but it also ignores the increase of the attack surface caused by offloading.

Hong et al. [103] focus specifically on the case of offloading over multiple hops in a network. They formulate the problem in game theory terms, so as to identify optimal strategies for offloading when taking into account computation time and energy consumption. This is an excellent approach because it offers self-configuration of each device's offloading strategy in a manner of achieving a Nash equilibrium, but again does not include resilience or security of the overall system as part of its quality of service optimisation.

9.3 Overview of innovative commercial products released since August 2018

In terms of commercial offerings, there are several new products or improvements of existing ones, but unlike C4IIoT's approach, their decision mechanisms are still taken manually by the user or automatically but based on predefined parameters.

In Cisco's Kinetic Edge & Fog Processing Module [104], the decision on where the processing will happen is still based on the system administrator choosing and applying the rules in advance. The rules can change at any time through human intervention. CISCO have recently updated their offering with a Gateway Management Module (GMM) [105], which is geared towards security management, yet does not take into account security in its offloading decisions.

The Intel IoT Gateway's capabilities were enhanced with the use of the Azure IoT Edge product [106], allowing it to offload computation with the use of software containers running locally on a smart edge device. It can be used to run mainly AI applications at the edge by keeping the most computation-intensive aspects of the application (e.g. the training) at the cloud.

Amazon has introduced its AWS IoT Greengrass [107] gateway, which some tasks to run on the IoT devices and others at the cloud. In practice, IoT devices can act locally on the data they generate, while the cloud does management, analytics, and durable storage. The compatible devices are able to run functions of AWS Lambda, which is an event-driven computing platform, so as to build serverless applications. However, the decision for offloading AWS IoT Greengrass is static and made by its user.

Google has also extended its cloud services to IoT gateways and edge devices with its Edge TPU [108] and Cloud IoT Edge products [109]. These also allow machine learning applications to run on edge devices using pre-trained models, but again the decision is mainly left to the user. [96]

10 Forensics visualization

10.1 Proposal's stated offering beyond August 2018 SotA

Digital forensics, as other areas where large numbers of unstructured data need to be collected and analysed, is based on data that must be stored and organized in such a way to allow patterns, relationships, etc., to be identified and explored in an effective manner. In most cases it is the linkages between the data that convey information and not the data itself. Hence, an effective visualisation tool should make the discovery, creation and presentation of these relationships as easy and as natural for the investigator as possible. Although a very powerful tool, machine driven analysis is not a silver bullet and its limitations can be leveraged by malcreants to escape detection. Therefore, at the end of the day, human input is required to complete the analysis. In this context, visualization plays a key role in augmenting the human analysis and providing context to the investigator. Especially in rapidly changing situations, the ability to focus on the important, rather than the mundane, is key to establishing situational awareness

The Forensics Visualization Toolkit (FVT) by AEGIS brings two innovative aspects to exceed current approaches:

(a) Timeline analysis, which provides the ability to “travelback in time” and compare the current situation with similar events that occurred in the past. This allows the new data to be compared against patterns encountered before. In this way, forensic investigators are able to identify malcreants from their M.O., and also create response strategies to be selected based on past knowledge of successful and unsuccessful outcomes;

(b) The ability to adapt the display of information based on previously encountered situations. For example, if the investigator has created a specific “view” consisting of multiple data sources and presentation modes (e.g. specific relationships to group and associate data items) to deal with a specific incident in the past, this view can be saved and reused, either manually or automatically, to present data associated with a new case, or an incident that is currently playing out.

Combined, these two aspects of FVT, allow the investigator to quickly gain a solid understanding of an event and benefit from existing knowledge gained from past interactions so as to identify the root cause of incidents and speed up the initiation of proper incident response actions.

10.2 Overview of research advances since August 2018

The explosive growth of smart objects and their dependency on wireless technologies for communication increases the vulnerability of Internet of Things (IoT) to cyberattacks and thus many researchers focus on this area. Digital Forensics Visualisation is a main domain of scientific research. In terms of research advances, we can point out interesting scientific publications such as:

- *scSVA: an interactive tool for big data visualization and exploration in single-cell omics* [110] The scSVA (single-cell Scalable Visualization and Analytics), is a lightweight R package for interactive two and three-dimensional visualization and exploration of massive single-cell omics data. Building in part of methods originally developed for astronomy datasets, scSVA is memory efficient for more than hundreds of millions of cells, can be run locally or in a cloud, and generates high-quality figures. In particular, we introduce a

numerically efficient method for single-cell data embedding in 3D which combines an optimized implementation of diffusion maps with a 3D force-directed layout, enabling generation of 3D data visualizations at the scale of a million cells.

- *Investigating Visualisation Techniques for Rapid Triage of Digital Forensic Evidence* [111]. A study that investigates the feasibility of a tool that allows digital forensics (DF) investigators to efficiently triage device datasets during the collection phase of an investigation. This tool utilises data visualisation techniques to display images found in near real-time to the end user. Findings indicate that participants were able to accurately identify contraband material whilst using this tool, however, classification accuracy dropped slightly with larger datasets. Combined with participant feedback, the results show that the proposed triage method is indeed feasible, and this tool provides a solid foundation for the continuation of further work.
- *Immersion or Diversion: Does Virtual Reality Make Data Visualisation More Effective?* [112]. When dealing with complex data, good visualisation tools are integral in efficiently inspecting trends and correlations. Virtual reality provides an opportunity to create a more immersive environment for users, utilising spatial awareness and depth perception. In this paper we investigate the effectiveness and usability of different systems using virtual reality and gesture control to visualise complex weather data. We propose three novel variants of our weather visualisation: a desktop-based application using an Xbox One Controller interface, and two virtual reality based applications, using an Xbox One Controller interface and a Leap Motion Interface. Our evaluation suggests that the Xbox One controller in combination with the VR display is most effective. Drawbacks of the gesture interface are its limited precision and reliability.
- *Digital Forensic Readiness for Financial Network* [113] proposed IP (Internet Protocol) traceback and visualization techniques for better digital forensics.

10.3 Overview of innovative commercial products released since August 2018

We are not aware of commercial products that were released since August 2018.

11 References

- [1] W. Report. [Online]. Available: http://www3.weforum.org/docs/WEF_Global_Risks_Report_2019.pdf.
- [2] ENISA. [Online]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018/at_download/fullReport.
- [3] EU. [Online]. Available: https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en.
- [4] [Online]. Available: <https://meltdownattack.com/>.
- [5] HPE, "The World's Most Secure Industry Standard Servers & How HPE Built the Silicon Root of Trust," [Online]. Available: <https://www.hpe.com/h22228/video-gallery/us/en/products/606a64ae-6d54-4c93-839f-b31b8e1903f8/the-worlds-most-secure-industry-standard-servers--how-hpe-built-the-silicon-root-of-trust/video/?lang=en-US>.
- [6] J. Kampeas, A. Cohen e O. Gurewitz, "Traffic Classification Based on Zero-Length Packets," *IEEE Transactions on Network and Service Management*, 2018.
- [7] M. A. I. M. Aminuddin, Z. F. Zaaba, M. K. M. Singh e D. S. M. Singh, "A Survey on Tor Encrypted Traffic Monitoring," *IJACSA*, 2018.
- [8] X. Zeng, X. Chen, G. Shao, T. He, Z. Han, Y. Wen e Q. Wang, "Flow Context and Host Behavior Based Shadowsocks's Traffic Identification," *IEEE Access*, 2019.
- [9] F. Pacheco, E. Exposito, M. Gineste, C. Baudoin e a. J. Aguilar, "Towards the Deployment of Machine Learning Solutions in Network Traffic Classification: A Systematic Survey," *Communications Surveys & Tutorials*, 2019.
- [10] P.-O. Brissaud, J. François, I. Chrisment, T. Cholez e O. Bettan, "Transparent and Service-Agnostic Monitoring of Encrypted Web Traffic," *IEEE Transactions on Network and Service Management*, 2019.
- [11] P. Zhou, W. Liu, L. Fei, S. Lu, F. Qin, Y. Zhou, S. Midkiff e J. Torrellas, "AccMon: Automatically detecting memory-related bugs via program counter-based invariants," *Proceedings of the 37th annual IEEE/ACM International Symposium on Microarchitecture*, pp. 269-280, 2004.
- [12] F. Long e M. Rinard, "Automatic patch generation by learning correct code," *ACM SIGPLAN Notices*, pp. 289-312, 2016.
- [13] G. Koschorreck, "Automated audit of compliance and security controls," em *Proceedings of the 2011 Sixth International Conference on IT Security Incident Management and IT Forensics*, 2011.
- [14] T. C. Chieu, S. Dutta, A. Gupta, A. McKay, B. Prysock, R. Ramaratnam, A. A. Shaikh, M. Signh, C. Tank and M. Viswanathan, "Automated Validation of Configuration and Compliance in Cloud Servers". Patent US Patent App. 13/419,591, September 2013.

- [15] A. Bolgert, R. Kalyanaraman, R. M. Forlenza and R. Cohen, "Supporting compliance in a cloud environment". Patent US Patent 9,110,976, Aug 2015.
- [16] S. Parthasarathy, S. Field, M. Goertzel, D. Kays, J. Dadzie and E. Reus, "Regulatory compliance across diverse entities". Patent US Patent App. 13/309,510, Jun 2013.
- [17] F. Doelitzscher, C. Reich, M. Knahl, A. Passfall and N. Clarke, "An agent based business aware incident detection system for cloud environments," *Journal of Cloud Computing: Advances, Systems and Applications*, vol. 1, no. SpringerOpen, p. 9, 2012.
- [18] F. Doelitzscher, C. Reich, M. Knahl and N. Clarke, "Understanding cloud audits," in *Privacy and security for cloud computing*, Springer, 2013, pp. 125--163.
- [19] C. Wang, S. S. Chow, Q. Wang, K. Ren and W. Lou, "Privacy-preserving public auditing for secure cloud storage," *IEEE transactions on computers*, vol. 62, no. IEEE, pp. 362--375, 2011.
- [20] C. Wang, Q. Wang, K. Ren and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in *2010 proceedings ieee infocom*, 2010.
- [21] B. Farinier, S. Bardin, R. Bonichon and M.-L. Potet, "Model generation for quantified formulas: A taint-based approach," in *International Conference on Computer Aided Verification*, 2018.
- [22] M. Anisetti, R. Asal, C. A. Ardagna, L. Comi, E. Damiani and F. Gaudenzi, "A Knowledge-Based IoT Security Checker," in *Euro-Par Workshops*, 2018.
- [23] M. Anisetti, C. A. Ardagna, F. Audenzi, E. Amiani, N. Diomede and P. Tufarolo, "Moon Cloud: A Cloud Platform for ICT Security Governance," *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1--7, 2018.
- [24] M. J. Farooq and Q. Zhu, "Modeling, analysis, and mitigation of dynamic botnet formation in wireless iot networks," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. IEEE, pp. 2412--2426, 2019.
- [25] M. Khodjaeva, M. Obaidat and D. Salane, "Mitigating Threats and Vulnerabilities of RFID in IoT Through Outsourcing Computations for Public Key Cryptography," in *Security, Privacy and Trust in the IoT Environment*, Springer, 2019, pp. 39--60.
- [26] I. Nadir, Z. Ahmad, H. Mahmood, G. A. Shah, F. Shahzad, M. Umair, H. Khan and U. Gulzar, "An Auditing Framework for Vulnerability Analysis of IoT System," in *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, 2019.
- [27] G. George and S. M. Thampi, "Vulnerability-based risk assessment and mitigation strategies for edge devices in the Internet of Things," *Pervasive and Mobile Computing*, p. 101068, 2019.
- [28] R. Negi, P. Kumar, S. Ghosh and S. K. Shukla, "Vulnerability Assessment and Mitigation for Industrial Critical Infrastructures with Cyber Physical Test Bed.," *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, pp. 145--152, 2019.
- [29] P. Speicher, M. Steinmetz, J. Hoffmann, M. Backes and R. Kunnemann, "Towards automated network mitigation analysis," *Proceedings of the 34th ACM/SIGAPP*

Symposium on Applied Computing, pp. 1971--1978, 2019.

- [30] S. Shafeeq, M. Alam and A. Khan, "Privacy aware decentralized access control system," *Future Generation Computer Systems*, vol. 101, p. 420–433, 2019.
- [31] A. D. Dwivedi, G. Srivastava, S. Dhar and R. Singh, "A Decentralized Privacy-Preserving Healthcare Blockchain for IoT," *Sensors (Basel)*, 2019.
- [32] "DPCM at IBM's website," [Online]. Available: <https://www.research.ibm.com/haifa/projects/imt/consent/index.shtml>. [Acesso em 2019].
- [33] "CPChain company website," [Online]. Available: <https://cpchain.io/>. [Acesso em 2019].
- [34] "CPChain White Paper 2.0," June 2019. [Online]. Available: https://www.cpchain.io/download/CPChain_Whitepaper_English.pdf/. [Acesso em 2019].
- [35] D. Lemos, G. D. R., C. Ghezzi and H. Giese, Eds., *Software engineering for self-adaptive systems 3: Assurances*, Springer, 2013.
- [36] C. Krupitzer, M. Breitbach, M. Roth, S. Van Syckel, G. Schiele and C. Becker, "A survey on engineering approaches to self-adaptive systems," *Pervasive and Mobile Computing Journal*, vol. 17, no. B, February 2018.
- [37] P. Lalanda, J. McCann and A. Diaconescu, *Autonomic Computing: principles, design and implementation*, Springer, Ed., 2013.
- [38] K. Pohl, G. Böckle and F. Van der Linden, *Software Product Line Engineering: Foundations, Principles and Techniques*, Springer, 2005.
- [39] D. Beuche, "Using pure:variants across the product line lifecycle," in *Proceedings of the 20th International Systems and Software Product Line Conference*, Beijing, China, 2016.
- [40] C. Krueger and P. Clements, "Systems and software product line engineering with BigLever software Gears," in *Proceedings of the 17th International Software Product Line Conference co-located workshops*, Tokyo, Japan, 2013.
- [41] T. Thum, C. Kästner, F. Benduhn, J. Meinicke, G. Saake and T. Leich, "FeatureIDE: an extensible framework for feature-oriented software development," vol. 79, no. 1, January 2014.
- [42] R. Capilla, J. Bosch, P. Trinidad, A. Ruiz-Cortés and M. Hinchey, "An overview of Dynamic Software Product Line architectures and techniques: observations from research and industry," *Journal of Systems and Software (SoSym)*, vol. 91, May 2014.
- [43] M. Brambilla, J. Cabot and M. Wimmer, *Model-Driven Software Engineering in Practice*, Morgan & Claypool, 2017.
- [44] N. Bencomo, R. France, B. Cheng and U. Assmann, *Models@run.time: foundations, applications and roadmaps*, Springer, 2014.
- [45] R. Mazo, J. Muñoz, L. Rincón, C. Salinesi and G. Tamura, "VariaMos: an extensible tool for engineering dynamic product lines," in *Proceedings of the 19th International*

Software Product Line Conference (SPLC'15), Nashville, TN, USA, 2015.

- [46] J. Muñoz, G. Tamura, I. Raicu, R. Mazo and C. Salinesi, "REFAS: A PLE approach for simulation of self-adaptive systems requirements," in *Proceedings of the 19th International Software Product Line Conference (SPLC'15)*, Nashville, TN, USA, 2015.
- [47] J. Mauro, M. Nieke, C. Seidl and I. Yu, "Context-aware reconfiguration in evolving software product line," *Science of Computer Programming*, vol. 163, 2018.
- [48] L. De Moura and N. Björner, "Z3 an efficient SMT solver," in *Proceedings of the International conference on tools and algorithms for the construction and analysis of systems*, 2008.
- [49] P. Stuckey, K. Marriott and G. Tack, "MiniZinc Handbook Release 2.2.3," [Online]. Available: <https://www.minizinc.org/doc-2.2.3/en/MiniZinc%20Handbook.pdf>.
- [50] S. Krieter, T. Thiem and T. Leich, "Using dynamic software product lines to implement adaptive SGX-enable systems," in *Proceedings of the 13th International Workshop on Variability Modeling of Software-Intensive Systems*, Leuven, Belgium, 2019.
- [51] D. Correa, R. Mazo and G. Giraldo, "Extending FragOP domain reusable components to support product customization in the context of software product lines," in *Proceedings of the International Conference on Software and Systems Reuse*, Cincinnati, OH, USA, 2019.
- [52] M. Weckesser, R. Kluge, M. M. M. Pfannemüller, A. Schürr and C. Becker, "Optimal reconfiguration of dynamic software product lines based on performance-influence models," Göteborg, Sweden, 2018.
- [53] N. Siegmund, A. Grebhahn, S. Apel and C. Kästener, "Performance-influence models for highly configurable systems," in *Proceedings of the 10th Joint Meeting of the Foundations of Software Engineering*, Bergamo, Italy, 2015.
- [54] J. Horkoff, F. Basak Ademir, E. Cardoso, T. Li, A. Maté, E. Paja, M. Salnitri, L. Piras, J. Mylopoulos and P. Giorgini, "Goal-oriented requirements engineering: an extended systematic mapping study," *Requirements Engineering*, vol. 24, no. 2, 2019.
- [55] E. Lara, L. Aguilar, M. Sanchez and J. Garcia, *Adaptive Security Based on MAPE-K: A Survey*, Springer, Cham., 2019, pp. 157-183.
- [56] GOV.UK Department for Digital, Culture, Media & Sport, "Collection Secure by Design - The Government's Code of Practice for Consumer Internet of Things (IoT) Security for manufacturers, with guidance for consumers on smart devices at home.," 28 February 2019. [Online]. Available: <https://www.gov.uk/government/collections/secure-by-design>. [Acesso em 6 September 2019].
- [57] A. Shamsoshoara, A. Korenda, A. F. and S. Zeadally, "A Survey on Hardware-based Security Mechanisms for Internet of Things," 2019.
- [58] S. Upadhyay, S. Kumar, S. Dutta, A. Srivastava, A. Mondal and V. Kaundal, "A Comprehensive Review on the Issues Related to the Data Security of Internet of Things (IoT) Devices.," Springer.

- [59] J. Stankovic, T. Le, A. Hendawi and Y. Tian, "Hardware/Software Security Patches for Internet of Trillions of Things."
- [60] A. Miloslavskaya and N. Tolstoy, "Internet of Things: information security challenges and solutions," Springer, 2019.
- [61] D. Levshun, A. Chechulin, I. Kotenko and Y. Chevalier, "Design and Verification Methodology for Secure and Distributed Cyber-Physical Systems.," IEEE, 2019.
- [62] M. Barbeau, G. Carle, J. Garcia-Alfaro and V. Torra, "Next Generation Resilient Cyber-Physical Systems," 2019.
- [63] MULTOS, "Chip to Cloud Security for the Internet of Things (IoT)," [Online]. Available: https://www.multos.com/uploads/SolutionBrief_MULTOS_ChiptoCloud.pdf . [Acesso em 6 September 2019].
- [64] Trustonic , "IoT Developer Kit – Secure By Design," [Online]. Available: <https://www.trustonic.com/microchip/iot-developer-kit/> . [Acesso em 6 September 2019].
- [65] Microchip, "Security ICs Design Center," [Online]. Available: <https://www.microchip.com/design-centers/security-ics>. [Acesso em 6 September 2019].
- [66] ASPENCORE, "Enhanced Cybersecurity for the IIoT," [Online]. Available: https://www.infineon.com/dgdl/Infineon-Enhanced_Cybersecurity_for_the_IIoT_MUST_READs-ART-v01_00-EN.pdf?fileId=5546d46269e1c0190169fc52b4276ab1. [Acesso em 6 September 2019].
- [67] I. E. Commission, "Edge intelligence," 2017.
- [68] "Microsoft," [Online]. Available: <https://azure.microsoft.com/en-us/services/iot-edge/>.
- [69] J. Konecny, H. B. McMahan, F. X. Yu, A. T. Suresh and D. Bacon, "Federated Learning: Strategies for Improving Communication Efficiency," *Computer Science - Machine Learning*, October 2017.
- [70] D. Vukobratovic, D. Jakovetic, V. Skachek, D. Bajovic, D. Sejdinovic, G. K. Kurt, C. Hollanti and I. Fischer, "A Reconfigurable Knowledge Acquisition Architecture for Future 5G IoT," 2019.
- [71] M. Cosovic, A. Tsitsimelis, D. Vukobratovic, J. Matamoros and C. A. Haro, "5G Mobile Cellular Networks: Enabling Distributed State Estimation for Smart Grid," *IEEE Communication Magazine*, Vol. 55, No. 10,, pp. 62-69, October 2017.
- [72] J. Pan and Z. Yang, "Cybersecurity Challenges and Opportunities in the New "Edge".
- [73] H. B. Review, "Most of AI's business uses will be in two areas," March 2019. [Online]. Available: <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/most-of-ais-business-uses-will-be-in-two-areas>.
- [74] M. Yazici, S. Basurra and M. Gaber, "'Edge Machine Learning: Enabling Smart Internet of Things Applications'," *Big Data and Cognitive Computing*, Sept 2018.
- [75] M. Yazici, S. Basurra e M. Gaber, "Edge Machine Learning: Enabling Smart Internet of

- Things Applications,” *Big Data and Cognitive Computing*, 3 Sept 2018.
- [76] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang and K. Huang, "Towards an Intelligent Edge: Wireless Communication Meets Machine Learning," Sept 2018.
 - [77] Wang et al., “Adaptive Federated Learning in Resource Constrained Edge Computing Systems,” p. doi: 10.1109/JSAC.2019.2904348, 2019.
 - [78] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo and J. Zhang, "Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing," May 2019.
 - [79] Y. He, J. Lin, Z. Liu, H. Wang, L.-J. Li and S. Han, "Amc: Automl for model compression and acceleration on mobile devices," *European Conference on Computer Vision. Springe*, p. 815–832., 2018.
 - [80] B. Zoph and Q. V. Le, "Neural architecture search with reinforcement learning," 2019.
 - [81] M. e. al., “Machine Learning at the Network Edge: A Survey,” p. arXiv:1908.00080, 2019.
 - [82] H. e. al., “Machine Learning in IoT Security: Current Solutions and Future Challenges,” p. arXiv:1904.05735, 2019.
 - [83] N. e. al., “DIoT: A Federated Self-learning Anomaly Detection System for IoT,” p. arXiv:1804.07474, 2019.
 - [84] F. & Saad, “Generative Adversarial Networks for Distributed Intrusion Detection in the Internet of Things,” p. arXiv:1906.00567, 2019.
 - [85] M. e. al., “N-BaIoT: Network-based Detection of IoT Botnet Attacks Using Deep Autoencoders,” p. arXiv:1805.03409, 2018.
 - [86] B. e. al., “IoTDS: A One-Class Classification Approach to Detect Botnets in Internet of Things Devices,” p. doi: 10.3390/s19143188, 2019.
 - [87] R. & Park, “Semi-supervised learning based distributed attack detection framework for IoT,” p. doi: 10.1016/j.asoc.2018.05.049, 2018.
 - [88] L. e. al., “ AI-Based Two-Stage Intrusion Detection for Software Defined IoT Networks,” vol. doi: 10.1109/JIOT.2018.2883344, 2018.
 - [89] D. e. al., “Machine Learning DDoS Detection for Consumer Internet of Things Devices,” vol. doi: 10.1109/SPW.2018.00013, 2018.
 - [90] P. e. al., “A deep recurrent neural network based approach for Internet of Things malware threat hunting,” vol. doi: 10.1016/j.future.2018.03.007, 2018.
 - [91] “5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies,” Gartner, August 2018. [Online]. Available: <https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/>.
 - [92] “Best Machine Learning Tools,” [Online]. Available: <https://www.trustradius.com/machine-learning>.
 - [93] “Top Open Source Tools for Deep Learning,” 2019. [Online]. Available: <https://www.rtinsights.com/top-deep-learning-tools/>.

- [94] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, no. 5, 2016.
- [95] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher and V. Young, "Mobile edge computing—A key technology towards 5G," *ETSI white paper*, vol. 11, no. 1, 2015.
- [96] M. Aazam, S. Zeadally and K. Harras, "Deploying fog computing in industrial Internet of Things and Industry 4.0," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 10, pp. 4674-4682, 2018.
- [97] L. Thames and D. Schaefer, *Cybersecurity for industry 4.0.*, New York: Springer, 2017.
- [98] X. Lyu, H. Tian, L. Jiang, A. Vinel, S. Maharjan, S. Gjessing and Y. Zhang, "Selective offloading in mobile edge computing for the green Internet of Things," *IEEE Network*, 2018.
- [99] M. G. R. Alam, M. M. Hassan, M. Z. Uddin, A. Almogren and G. Fortino, "Autonomic computation offloading in mobile edge for IoT applications," *Future Generation Computer Systems*, 2019.
- [100] A. Alelaiwi, "An efficient method of computation offloading in an edge cloud platform," *Journal of Parallel and Distributed Computing*, 2019.
- [101] X. Li, J. Wan, H. Dai, M. Imran, M. Xia and A. Celesti, "A hybrid computing solution and resource scheduling strategy for edge computing in smart manufacturing," *IEEE Transactions on Industrial Informatics*, 2019.
- [102] S. Chen, Y. Zheng, K. Wang and W. Lu, "Delay Guaranteed Energy-Efficient Computation Offloading for Industrial IoT in Fog Computing," in *IEEE International Conference on Communications (ICC)*.
- [103] Z. Hong, W. Chen, H. Huang, S. Guo and Z. Zheng, "Multi-hop Cooperative Computation Offloading for Industrial IoT-Edge-Cloud Computing Environments," *IEEE Transactions on Parallel and Distributed Systems*, 2019.
- [104] CISCO, *Cisco Kinetic Edge & Fog Processing Module*, 2018.
- [105] CISCO, *Cisco Kinetic Gateway Management Module (GMM)*, 2019.
- [106] "<https://azure.microsoft.com/en-gb/services/iot-edge/>," [Online].
- [107] "<https://aws.amazon.com/greengrass/>," [Online].
- [108] "<https://cloud.google.com/edge-tpu/>," [Online].
- [109] "<https://cloud.google.com/blog/products/gcp/bringing-intelligence-edge-cloud-iot>," [Online].
- [110] M. Tabaka, J. Gould and A. Regev, "scSVA: an interactive tool for big data visualization and exploration in single-cell omics," 2019. [Online]. Available: <https://www.biorxiv.org/content/early/2019/01/06/512582.full.pdf>.
- [111] G. Hales and E. Bayne, "Investigating Visualisation Techniques for Rapid Triage of Digital Forensic Evidence," in *Proceedings of the 21st International Conference on Human-Computer Interaction*, Orlando, FL, USA, 2019.

- [112] B. Andersen, A. Davis, G. Weber and B. Wünsche, "Immersion or Diversion: Does Virtual Reality Make Data Visualisation More Effective?," in *Proceedings of the 18th International Conference on Electronics, Information and Communications*, Auckland, New Zealand, 2019.
- [113] S. Kwon, J. Jeong and T. Shon, "Digital Forensic Readiness for Financial Network," in *Proceedings of the International Conference on Platform Technology and Service*, Jeju, South Korea, 2019.